Visual sampling processes

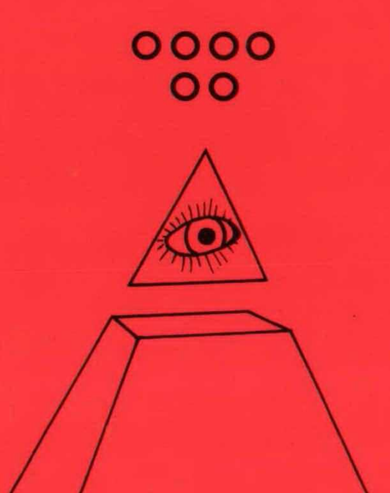
Senders, John Warren

Publication date:
1983

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Senders, J. W. (1983). *Visual sampling processes*. [Doctoral Thesis, Tilburg University]. [s.n.].

VISUAL SCANNING PROCESSES



JOHN W. SENDERS

VISUAL SAMPLING PROCESSES

PROMOTOR : PROF. DR. A.F. SANDERS

VISUAL SAMPLING PROCESSES

Proefschrift

ter verkrijging van de graad van doctor in de sociale wetenschappen
aan de Katholieke Hogeschool Tilburg op gezag van de rector magnificus,
prof. dr. R.A. de Moor, in het openbaar te verdedigen ten overstaan van een
door het college van decanen aangewezen commissie in de aula van de
Hogeschool op vrijdag 2 september 1983 te 16.15 uur

door

JOHN WARREN SENDERS
geboren te Cambridge (Mass.) V.S.

1983

drukkerij Neo-Print Soest

STELLINGEN

I

The visual fixations of instrument monitors are related to the bandwidths of the signals which drive the instruments in a direct way. Increasing the bandwidth leads to an increase in fixation frequency.

II

The behaviour of the monitor is describable as that of a single channel queue. The statistics of the queueing process are derived from the sampling strategy adopted by the observer.

III

Each sampling strategy gives rise to a mathematical model of the statistics of fixational behaviour. Some of these models are related to the statistics of the signals; others are related to the actual values of the variables which the monitor observes.

IV

The informational-theoretic conceptual model of the human operator is a robust one. It permits excellent qualitative prediction of behaviour. The notion that Information Theory is not appropriate to the description of human behaviour iś quite incorrect.

V

Human beings are uncertainty-reducing systems. This reduction in uncertainty is accomplished by information-seeking and finding. Eye movement behaviour is explained by this simple yet powerful pair of statements.

VI

Psychology abandons its conceptual schemes too early, usually because the schemes do not explain everything. Yet to be able to explain much is often the test of a good model.

VII

Psychology has concentrated on correct performance and used error statistics to measure the level of performance. Yet it is the errors which can destroy us. These should be the new target of psychological research.

VIII

The application of mathematical models derived from engineering to the problems of psychology has had more beneficial effect than have the application of mathematical models derived from psychology to the problems of engineering.

IX

Three Mile Island has done more for the field of Engineering Psychology than all the special pleadings of its promoters in the past.

X

It is a moot point as to whether the eye moves where the mind wants it to go or the mind sees what the eye wants to show.

XI

The use of voluntary occlusion of vision permits the externalization of the elusive process of attention. The study of automobile driving without looking demonstrated that the technique gives a direct measure of the demand placed on the driver by the task.

XII

The fundamental problem of errors and accidents is whether errors are caused or uncaused. If they are uncaused, they are mental "acts of God" and people should not be blamed for them. The implications for the concept of responsibility are enormous.

XIII

The question of responsibility is simple : You are responsible for your errors; I am not responsible for mine.

XIV

The PhD degree in recent years has often become an indication of promise. Better to receive it upon retirement when the promise has been fulfilled.

Stellingen bij : Visual Scanning Processes

J. W. Senders Tilburg, 2 september, 1983

ACKNOWLEDGEMENTS

I am especially grateful to Prof. Andries Sanders who proposed to me that I might undertake to complete a Ph.D. by dis-assembling and re-assembling scientific material that I had done over a span of more than 30 years and presenting it as a dissertation. I have corrected many of my old errors; reanalyzed many of my old data; discovered new and better ways of looking at my old problems; arrived at new answers to some of them and rejected some of the old answers to others. All in all, I have learned much from the process.

I am deeply appreciative of the time and effort put in by Prof. J.E. Rijnsdorp, Prof. W.A. Wagenaar and Dr. Henry Beller for the careful reading, and criticism they have done. Errors in logic and mathematics have been found and corrected by that process. Whatever errors remain are purely my own responsibility; perhaps I was too persuasive for their critical eyes.

Michael Venturino did the graphics and made it possible for me to meet deadlines. He did the good figures. The bad ones I did myself.

Ann Crichton-Harris, through her consistently moral strength and urgency, made sure that I did what I had set out to do. Sine uxore nihil !

MORE ON THE EYE

The eye is both a servomechanism and a mécanisme de cerveau.
And sometimes it does its own thing and somtimes it goes
        Where the brain wants it to go.

The eyes are the window to the mind and the mind's window
        To the scene
So that one is never quite sure whether it's the world or
        The mind that makes the eye shift to where it's going
        From where it's been.
You can watch the eyes and catch the thought
While it's so hot that even the mind hasn't had it yet.

With a mind of its own the eye looks at the place best calculated
        To let the mind's eye see what the mind wants to see;
And then all the world rushes in to be reduced
To common sense and percept before the next saccade is loosed.

J.W. Senders
August, 1980

CONTENTS

CHAPTER 1 - INTRODUCTION

1. Eye Movements and Scanning.

Vision for human beings, like audition for bats, is surely the principal channel through which information is acquired about the outside world, The physical and functional anatomy of the human eye, however, makes it impossible for information to be gathered equally efficiently from all parts of the surrounding world. Not only is the entrance angle of the eye limited to 180 degrees more or less, but there is also a reduction in resolution for objects or stimuli more than a few degrees from the line of sight. This reduction stems from three sources : 1) the retina has its greatest packing density of receptor cells around the center of the visual field; 2) these central receptor cells are connected on a one-to-one basis to corresponding cells in the cortex. Further away from the center of the visual field, the cells tend to become interconnected at ganglion layers immediately adjacent to the receptive surfaces; 3) there are refractive errors which degrade visual resolution away from the center of the visual field. As a consequence one does not see very well more than two or three degrees away from the line of sight. A human being must move the eyes either by moving the head, or by moving the eyes within the head, in such a way as to bring into the central part of the visual field whatever is to be examined in detail.

The human eye is thus constantly in motion under normal conditions. Part of this motion is the result of voluntary effort on the part of the observer who deliberately moves the point of regard from one place in the visual field to another. Most of the movements, however, are involuntary, and the observer is quite unaware of a good part of them. There is an in-between category where in the observer is aware that the eyes are moving, but mistakenly assumes they are making one kind of movement when in fact another is going on.

Since the eye is the most powerful channel that the brain has for taking up information, it is important to know where the eye looks and why, how long it looks there and why, what is there to be seen and where it looks next and why, if one is to understand something of what is happening in the brain.

Eye movements can be sorted into the large and small, and into the fast and slow. The large fast movements are called saccades. Slow small movements are called drifts. Large slow movements are called pursuit movements. Finally, fast small movements are called flicks or microsaccades. Even when the eye is fixated voluntarily there is residual movement. A typical eye movement record of a "fixated" eye will show a succession of random appearing drifts and flicks with the general position of the eye held fairly steady. On the average, the position of the fixated eye remains constant within a very small area, perhaps less than one quarter degree in radius. Figure 1 shows a series of saccades made during reading (a) and another made during instrument panel watching (b). In both cases the eye moves at high velocity from one point of regard to some new point of regard. During reading, the succession of fixation points is organized roughly according to the printed information so that a succession of left to right saccades occurs, broken from time to time by a very large right to left saccade as the eye returns to the beginning of the next line. During instrument reading there is also a continuous series of saccades, but their direction is not apparently organized in any particular way except that certain patterns seem to recur on an aperiodic basis. Voluntarÿ eye movements reveal to an outside observer what it is that an inside observer, the "mind", desires to accomplisch and desires to obtain from the outside world. In other words, the voluntary movements of the eye present information in a very direct way about internal mental processes, in particular the process of visual attention (except when the eye mover is trying to deceive). There is a stabilizing system that protects the image on the retina against body motion so that if the eye is fixated and head motion occurs, the fixation is not disturbed. Somewhere between these completely automatic stabilizing activities of the eye and the "intentional" system earlier described, which allows the observer to redirect the eyes on a voluntary basis, is a system which I will call the "eyes'mind". The eyes'mind directs the larger part of all the involuntary saccades. Many of these go on without awareness on the observer's part. The eyes'mind, in a sense, serves as a gatekeeper between the changing external world of visual stimuli and the internal representation of that stable visual world, the mind's eye. The eye movements and the selection of points to which the eye is redirected stem from an evaluation by this eyes' mind of the needs of the information processing system required for the maintenance of the image in the mind's eye. In brief, then, the eyes mind tells the eye where to look in order that the mind's eye will see what it wants to see.

2. The Pilot Eye Movement Studies.

The classic data on visual sampling of displays in an aircraft are those gathered by Milton, Jones, Fitts et al. and described in a series of reports published as technical reports of the laboratories at Wright Patterson Air Force Base, Ohio. The work started in 1949 and continued over many years (1) Indeed, further studies on more modern aircraft were carried on until 1966. In a summary report published in 1950, Fitts stated : "It is reasonable to assume that *frequency* of eye fixations is an indication of the relative *importance* of that instrument. The *length* of fixations, on the contrary, may more properly be considered as an indication of the relative *difficulty* of checking and interpreting particular instruments. A *pattern* of eye movements, i.e., the Link Values between instruments, is a direct indication of the goodness of different panel *arrangements.* "(All emphasis from original). Fitts further commented : "Information about how pilots use their eyes while flying on instruments in fundamental to a basic understanding of the functions served by aircraft instruments and to simplification of the psychological processes that occur while a pilot is controlling an aircraft's altitude, location, and rates of movements in three-dimensional space. If we know where a pilot is looking, we do not necessarily know what he is thinking, but we know something of what he is thinking about". The goals of Fitts and his colleagues were to answer questions like these : Do pilots observe more than one instrument at a glance, or do they always look directly at specific instruments, one at a time ? Are some instruments checked much more often than others ? Do some instruments require much longer to check than others? "It was believed that answers to these questions would help to solve problems that arise in designing instruments that are easy to read and in arranging these instruments into functional groupings on the instrument panel. "The underlying theme of the conclusions is about as follows : Those instruments at which a pilot looks often should be located in the central area of the instrument panel; Those instruments which are looked at for a long time are either very important and should be centrally located and possibly enlarged, or are difficult to read and should be redesigned; those pairs of instruments which are looked at often in succession (high link values) should be located close to one another.

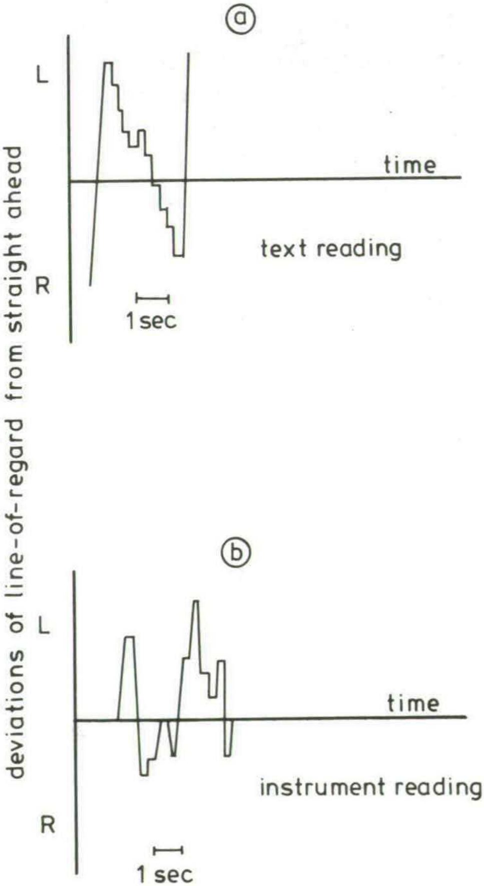


Figure 1 : Change of Line-of-Regard as a Function of Time (Typical)

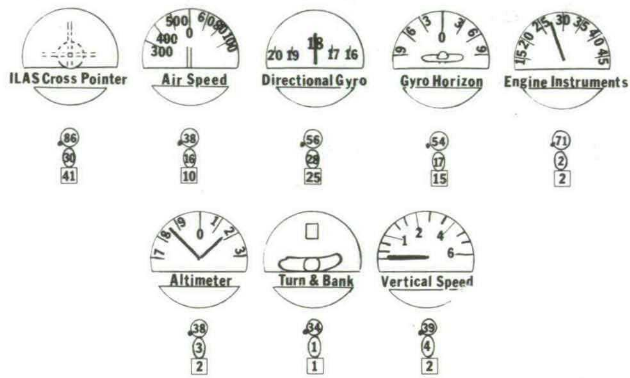
Thus, it was reasoned that if one measured the actual eye fixation behaviour of pilots engaged in various flight maneuvers under various conditions, one could find the best way to arrange the instruments on the panel. The approach was, therefore, completely empirical : the behaviour was to be examined and the design was to be based on the behaviour. Whether the design on which the data were based affected the behaviour was not in general questioned. Nor was any theory put forward to relate the looking behaviour to what the pilot actually saw on the instrument looked at.

The result of the studies was a great set of data about frequencies, durations, and sequences of eye fixations on instruments of various names like altimeter, gyro horizon, cross-pointer and the like. The data for each flight condition, each maneuver, and each aircraft were summarized in figures an example of which is shown in Figure 2. It should be noted that the link values shown are the aggregate of the transitions in both directions between the instruments named. Thus, if there were important asymmetries, they are not presented in the tables or the figures.

The numbers below each instrument show the frequency of fixation, the duration of fixation, the percentage of total time spent on that instrument; the number on the arc connecting a pair of instruments is the percentage of all transitions made by the subjects that were made between those instruments. For example, the Directional Gyro was looked at 25 times per minute, for an average duration of 58 seconds, and 25 percent of the total time was spent looking at it. The transitions between the Directional Gyro and the Gyro Horizon were 15 percent of all the transitions made. These results were valuable in establishing good instrument panel layouts. The modern arrangements of aircraft instruments stem from these studies.

What the studies could not do was to enable the investigator to establish rules that would permit generalization of the data to other aircraft, other flight conditions, other maneuvers, or new instruments. Even less could the results be used to calculate the instrument problem solution for some new and untried system like a space vehicle or a nuclear power plant control room neither of which had been conceived at that time. This was so

Length of Eye Fixations and Number of Fixation of Air-Craft Instruments During ILAS Approach

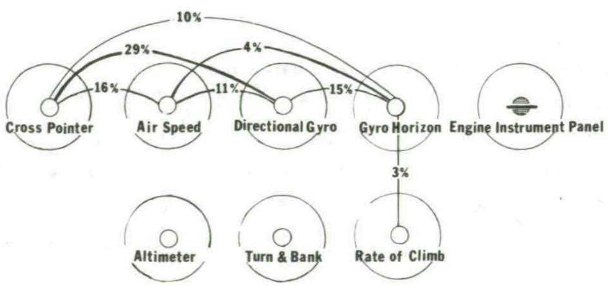


Length of Fixation Cycle ○ seconds

Number of Fixations per Minute 0 number

% Time Spent on Each Instrument □ %

Eye Movement Link Values Between Aircraft Instruments Instrument Landing Approach System (ILAS)



Link values based on 40 Pilots

Values less than 2% omitted

because no information was gathered about what the instruments said to the pilot. Only one side of the transaction was recorded and, therefore, no connection could be made relating what the instrument said to the behavior of the person who looked at the instrument. No general theory of instrument scanning or of any kind of monitoring behavior could be erected. Since no normative predictions of behaviour could be made, there was no way other than by repeated empirical study that the ambiguities could be resolved. Each new aircraft and each new instrument had to be measured anew.

Obviously, if there was to be a generalizable theory of instrument watching, the nature of the signal presented on the instrument had to be considered.

3. *The Modern Problems.*

A thoroughly modern task is that of monitoring a large number of instruments that present information about the state of a large system. A nuclear power station control room presents such a task. Control room operators work in an environment that may contain upwards of 2500 instruments and controls. Each of these instruments presents information about the state of a variable, and the controls allow modification of a variable. The task of the operator is to take in information about the state of the system and, when appropriate, alter system variables according to some rational procedure. The operators must move both their eyes and bodies about in the work space in order to be able to read the various instruments and to manipulate the various controls. Since, with only a few exceptions, the instruments not being fixated are outside the fixation field of the operator, such an instrument cannot be seen until it is actually fixated, and the choice of what to fixate or what to observe must arise from some process within the operator, Some mental process leads to the attentional act. The monitoring of the many dials of any complex machine must be done by looking first at one and then at another. Even if some of these instruments have more than one kind of information displayed, most are single variable, and most are designed for foveal viewing. This last fact arises partly because people do fixate on things they want to see and partly because designers tend to look at the things that they are designing and therefore design them for foveal viewing. Thus successive fixation is necessary. Since that is so it becomes of great interest, both theoretical and practical, to discover the laws which govern such successive attentive behaviour. The practical applications are straightforward : if one knew the laws governing such scanning behaviour, one could design better display panels, calculate work loads, and so on. The theoretical applications are also easy to see : if there are lawful relationships between overt attention shifts and the material which is obeserved, then these same laws form a basis for modeling and experimentation on the hypothesized covert attention shifts that may be going on as described by Posner. (2).

If one puts oneself in the position of the controller of a complex machine, an aircraft for example, one finds that the looking at a particular instrument is not always, or even often, the result of a conscious decision. The more practiced the operator the less likely is the looking a consequence of conscious decisions. However, when the eye is at a moment fixated on a particular instrument in the array of instruments, the information presented at that point is attended to. Thus the overt behaviour of eye movements appears to be driven by the anticipated attention and is not itself the attention. Eye movements are correlated with attention and are manifestation of attention*.

When sources of useful information are placed sufficiently apart in space, some kind of overt sampling behaviour must take place. The separation need not be very large since the visual acuity of the human eye falls off rapidly with angular deviation from the point of regard. Naturally, if displays of information were to be made large enough, there would be no absolute necessity for eye movements to take place. Nonetheless, unless suppressed by conscious effort, they will occur. In operational systems, the number of types and sources of information to be processed by the human operator is so large that the displays are virtually mandated to be legible only by foveal vision. In any event, most visual displays have been designed by designers who look at what they are drawing and therefore make displays suitable for foveal viewing.

It would be the rare display designer who designed while looking at a point 20 degrees to the left rather than directly at the point of the pen. For these and other reasons, human beings will make eye movements when they are confronted by many things to look at. In what follows, models will be constructed that relate the frequency and duration of eye fixations on any one information source to properties of the signal or message which appears on that information source; and that relate the sequences of eye fixations on a number of information sources to the properties of the entire ensemble of signals (or messages) displayed on the various sources.

* So called "head up" displays are sometimes used where even the loss of time incurred in making eye movements is intolerable. Naval carrier landing operations are one example.

In homely terms : why does a pilot, confronted with many instruments on an instrument panel, look at instrument *I* at time *T* for duration *D* following a previous look at instrument *I*, *X* seconds ago ? And why does that pilot look at instrument *I* more often after looking at instrument *J* than after looking at instrument *K* ? Explicitly : (to the pilot) - Why did you look at the altimeter just then for .48 seconds ? Why did you then look at the attitude indicator for .39 seconds? Pilots have a difficult time answering questions like these. Often pilots are not aware of the nature of their own eye movement patterns and are grossly in error about them, believing that they use regular left to right successive sweeps across the instrument panel when is fact they are essentially random fixators as will be shown.

4. The Present Work.

This work will attempt to establish lawful relationships between what human observers look at and what the instruments they look at tell them, in quantitative formal ways. The results of the analysis will be particularly useful in calculating the work loads imposed on human beings by complex monitoring and supervisory control tasks. It is not clear that those persons who design such important artifacts of modern life as nuclear power plant instrument control rooms have any secure basis for calculating how many human beings are required to operate a particular plant. There are obvious disadvantages to having too many people and obvious dangers to having too few. A theoretical basis for calculating how many people would be required to observe all the instruments would be at least one way of estimating the manning requirements of such systems. The importance of being able to do this can not be overestimated. Many modern systems involving large quantities of power and toxic materials are under the control of relatively few people. As a consequence, even though the reliabilities of hardware have increased dramatically with the introduction of digital control systems, the potential hazard associated with system failure has grown even more.

In brief, Chapter 2 will present a set of models of the fixation-generating mechanism of the human monitor. These models provide input data for a queue-theoretic analysis of scanning behaviour. Chapter 3 contains the results of a series of experiments on visual scanning and offers an evaluation of the various models. Chapter 4 considers the implications for behavioural theory and for application and discusses some unanswered research questions that are of importance for both theory and application. Chapter 5 is a concluding synthesis and summary.

There are a number of assumptions which have to be made. Foremost among these is that human beings behave like uncertainty-reducing machines. The method by which uncertainty is reduced in the seeking and finding of information. The various appendices are all directed toward the study of this aspect of behaviour. Appendix A attacks the problem of the distribution of visual attention over an array of information sources. The calculations based on the model of fixation duration show that the duration is directly related to the amount of information to be taken in or, mutatis mutandis, the amount of uncertainty which had accumulated since the last fixation on the same instrument. If, as there may well be, there is a difference between the subjective and the objective statistics of the signal, then the subjective uncertainty measure would govern rather than the objective information measure. Since, however, we also assume that all the observers are highly practiced and in a sense ideal, these are treated as if they are identical. The observer is further assumed to have perfect recall. There is no forgetting of what has been seen on the previous observations, in contrast to the model in Appendix B. The sampling behaviour is driven only by the need to have samples sufficiently frequent to allow reconstruction of the signal on the basis of the sample data.

Appendix B, the study of automobile driver information processing, introduces the notion that uncertainty arises both from without and within the human observer. As the driver passes over the road, new information comes into view and must be taken in when the driver looks at the road. In the intervals between looks at the road, the memory of what the driver has just seen fades; this in turn generates uncertainty about what in fact was seen on the occasion of the preceeding look. The driver knows less and less as time passes from one look to the next where the vehicle is with respect to obstacles and boundaries. The two uncertainties together make up the total uncertainty possessed by the driver on the occasion of the next look at the road. This model also fits the instrument scanner with respect to each of the instruments to be monitored. In the case of the driver, the assumption is made that there is a threshold of permissible uncertainty. When that threshold is reached, the driver looks. One of the models of the instrument monitor uses the same mechanism to calculate the intervals between fixation on the various instruments in an array.

Appendix C presents a conceptual model of workload based on sampling of elemental capacities or talents. The underlying assumption is that of a unitary attentional system that can deal with only one kind of activity at a time. The model is an analogue, therefore, of the visual monitoring process which so strongly depends on foveal vision.

CHAPTER 2 : A THEORY OF VISUAL SCANNING BEHAVIOUR

1. *Introduction*

An ever increasing number of tasks in many human endeavors involve visual monitoring of large numbers of visual displays of information : instruments. Although this trend is partly the result of the increase in the number of digital processors used in the control of processes, it had begun long before the digital computer became a commercial product. Particularly in military activities, but also in electrical power generation and chemical process control, the design of control systems for human use was based on human visual capacity and flexibility. The control room of a coal-burning power plant was and is a very complex array of instruments and controls. This is also true of the cockpits of aircraft whether from the 1940's or the 1980's.

There have been no good ways of assessing the load placed on the human system component or of determining how many human operators (H.O.) any particular system really needed for safe and efficient functioning. Control rooms and cockpits have been designed for specified numbers of operators but not on any analytical basis. As a result, some systems have been satisfactory and successful and others have not. The calculation of the reliability of complex man-machine systems has been hindered, if not made quite impossible, by this lack of analyticity.

One of the characteristics of the monitoring behaviour of a skilled operator is that most of what goes on is quite automatic. There is rarely a conscious decision to "look at the *z* pressure gauge". Instead the pressure gauge is looked at about when it should be in the midst of a sequence of looks at many other things. Thus far, the systems so monitored seem to do what they are supposed to do. It might be imagined that this is a matter of great good fortune and that sooner or later the lack of systematic conscious decision-making would lead to system failure. On the other hand, it is when conscious instead of automatic skilled behaviour goes on that the loads on the human operator become large and the plant unmanageable.

An unskilled operator placed in the job will find that the demands are very great and that proper management is impossible under many circumstances. Clearly the experience of the operator used to the task plays a role in the maintaining of system performance. Since much of what goes on in the instruments of a complex system is not completely deterministic, there is always *some* statistical variation; the experienced operator must have learned something about the statistics of the displayed signals given that H.O. is able to look where looking is needed rather than where it is not. Any theory or model of the behaviour must include this characteristic.

Control room operators and pilots attend to instruments. This is manifested by the fact that they fixate on the instruments. It is not too important from the practical, engineering point of view whether the eye movement and subsequent fixation *are* attention, are *related* to attention, or are unconnected with attention. If one could predict something about fixations, one would have solved some of the more vexing operational problems. From the point of view of the behavioural theorist, on the other hand, the question is of significance. Although attention can be directed by instruction, the process of automatic attending may well be different and have different laws governing it. In the context of this analysis there is no need to assume such differences; only the source of the command to shift the locus of attention is assumed to be different. In light of recent studies by Posner and his colleagues (2), it is reasonable to assume that the skilled operator has an adequate internal spatial map of the array of instruments being worked with, and knows where the various instruments are as well as what kinds of signals appear on them.

When the central representation of an instrument makes an "attentional demand", attention itself will be directed toward the external, real instrument and the eye movement will follow to bring the display into foveal vision. This is necessary for the reading of the information presented on the instrument, since they are uniformly designed for foveal viewing.

Thus, attention is not the eye movement itself but is the driving force which directs and motivates the eye movement.

The underlying assumption of the analyses to follow is that attention is itself generated, with respect to a source of information, by subjective or internal uncertainty about the state of some variable or constant aspect of the external world. The control room monitor redirects the gaze in order to reduce uncertainty to an acceptable level and reads the data presented on each instrument to some degree of precision which meets a criterion. The criterion might be specifically instructed or it might be the consequence of long exposure to the instrument.

Where does the uncertainty come from ? The monitor may forget what was seen on the previous observation. Since many observations of other instruments may have intervened, this is likely, given modern data and theory of short term memory. There will have been retroactive inhibition with respect to the instrument information gathered some moments before as a result of having gathered many other kinds of information since that time. Thus, even for a constant real world state, the monitor will from time to time make an observation. The other source of uncertainty arises from the statistics of the variables in the real world. Most instruments vary and the variation, which is the result of variation in the underlying processes, has determinable statistical characteristics. Some will have a large range and some a small one. Some will vary quickly and some slowly. The variation means that during the interval between observations of an instrument uncertainty must be generated in the monitor about what actually is displayed. If the monitor's internal representation of an instrument is not in accord with reality, then the amount of uncertainty may be different from some true value. This, however, is a matter of the experience of the operator. The total uncertainty at some time after an observation is the sum of these two independent quantities. The uncertainty may itself be about various aspects of the signals that drive the instruments. For example, one of the models will assume that the monitor is uncertain about the actual signal itself. H.O. takes on, therefore, the task of reconstructing the signal on the basis of the samples taken. At the other end of the spectrum of models is one based on the idea that detection of divergent events is the task. This monitor behaves quite differently from the first mentioned. Whatever the goal taken by the monitor, some strategy of looking will result.

Although in the real world of monitoring there are alarm lights and sounds which make imperative demands for attention, the models do not consider these. Instead, all the models deal only with steady-state, non-emergency condition. It is assumed that, when an emergency arises and an alarm operates, the behavioural rules change. Some quite different model is needed for that state of affairs.

What is the action of the increasing uncertainty in the observer ? If the uncertainty is below some implicit threshold, no observation will be made. When the uncertainty rises so that it exceeds that threshold, attention is demanded and an eye fixation results. Such an assumption implies that the observer will not look at anything if none of the instruments (the internal representations, of course) has an uncertainty in excess of the threshold. An alternative assumption is that the observer always looks at something and chooses to look at that instrument whose internal representation is maximally uncertain. There are other possibilities as well.

Although the principal orientation of the analysis is toward the multi-instrument task, the general notions can be applied to tasks in which information is provided globally and in which the locus of "interesting" information may be quite ill-defined. Automobile driving is one such task. The mathematizing which results in the model building process is different since it seeks to explain what happens between looks at the whole world rather than at a small segment of the world. Appendix B presents some results from analysis of that kind of task.

Still another application can model the scanning which takes place when observers are presented with pictorial displays of unchanging nature. A painting, for example, does not change , at least in the real world. It surely changes in memory in that the store about any segment has been interfered with by the looks at and uptake of information from other segments. There may also be changes in the criterion of knowledge of the picture such that the first appreciation of a painting may be of global execution; the last may be of detail of brush work. As the nature of the information to be read out changes, the uncertainty in the internal representation must grow.

For both the global varying task like driving and the global invariant task of picture-looking, there will be characteristics of the scene which have something to do with the distribution of attention, as represented by eye fixations, over the scene. The density distribution of detail in a picture, for example, should rationally have an influence on the density of attentive behaviour. In a more quantitative way, this aspect can be considered to be a function of the spatial bandwidth of various portions of the scene. The portions with the high bandwidths have the higher information densities and should, ceteris paribus, demand the more attention. There are , of course, difficulties with this formulation that result from the fact of "meaning", which tends to interfere with the easy description of human behaviour in mathematical ways. Naturally, in automobile driving, peripheral "streaming patterns" may be much more meaningful and therefore more important than the high frequency signals from the leaves on the trees!

For any task, there is a necessity that the skilled operators be considered to be working at the *margin*. When an operator is working at the margin, there is little or no room for behavioural choice. If there were performance would not be at the margin. The margin may be defined as the level and kind of behaviour that is dictated by the task. Literally, the human operator is controlled in form and function by the configuration of the task; human dynamics are strongly related to the task dynamics; human goals are strongly controlled by the goals of the sytem. In one sense, the human operator is as close to a robot as possible and there may be little psychology remaining. Of course, no human operator is long at the margin. It is only that H.O. approaches it from time to time when H.O. is able and is required to do so by task demands. No amount of skill completely forecloses the need for practice or defends against the loss of some competency with only a little disuse of that skill.

The theory and the attendant models, therefore, apply only to demanding tasks performed by experienced and skilled human beings. No novices need apply.

In summary, the internal representations in the monitor of the real spatially disposed information sources from time to time demand attention because H.O. is excessively uncertain about them. The attention shifts to the real world source of the needed information and "drags" the eye with it. The external manifestation of the internal attentional shift is the eye movement. For certain classes of signals the timing of the demands can be estimated. The intensity of the demand, which corresponds to the level of uncertainty in the observer, can be estimated on the basis of the statistical characteristics of the signals and the functional characteristics of the monitor. A general measure of the "attentional demand" made by a source of information on the monitor is the relative amount of time spent on that source of information.

Finally, there must be a meaningful task with well-defined goals and some reason for getting it done. Neither theory nor model attempts to take into account fatigue, motivation, or anything else of purely psychological nature. These monitors are perfectly motivated and never fatigued as well as being completely skilled.

2. The Queueing Model

The general theory of instrument monitoring is predicated on a queueing model of the scanning process as it unfolds in time. An observer with only one instrument to monitor, if no other demanding events occurred, would be able to look at that instrument whenever a demand arose. (Forget for convenience the apparent nature of human vigilance performance since we are dealing with robot-like human operators). No event which should have been observed will go unobserved, assuming that the strategy of the monitor is a good one; nor would there be any delay in making the observation since there would be nothing to intervene in the process of translating demand into action. On the other hand, if there were two or more instruments, each with its own characteristic rate of uncertainty generation, and, therefore, with its own distribution of demand in time, there is a probability that simultaneous demand will occur. The observer will be looking at one instrument when the uncertainty of the other exceeds its threshold and calls for attention. The analysis of attention can thus be approached as a problem in queueing theory when there are competing stimuli to be attended to.

It may be the case that an instrument will have to wait until others which have entered the queue earlier have been attended to before its turn comes. Naturally there is some probability that the event which the demand was based on, the event about which there was too much uncertainty, will occur during this wait. In that case, there may be considered to have been a failure of the man-machine system (in the practical case). The analysis of the reliability of complex man-machine systems may also be approached as a problem in queueing theory. Appendix D analyzes these problems.

Queueing theory is the mathematical analysis of systems involving one or more "service channels" and some number of "customers" who arrive at the service channels for service. Queues are a part of everyday life. We encounter them at the bank, at the post office at the ticket counter at the airport, at the supermarket checkout line. Sometimes there are many service channels open and at other times only one. The other queues, where each service channel has its own queue, always seem to move faster than one's own. The only solace is the thought that the observation is as true for the people in the other queues.

A service channel takes time to service the customers who arrive. Usually the time is not a constant. Instead it is a variable which has some distribution, sometimes simple and exponential, more often not. The intervals between customers ' arrivals at the queue of a service channel also form a distribution, sometimes simple and exponential, more often not. For any service channel it is clear that if the service time is short and the interval between customers long, the service channel is frequently not in use. On the other hand, when the service times are long and the intervals between customers short, there will frequently be a line of people waiting. When the interval between customers is actually shorter, on average, than the service time, on average the queue will grow without limit, and the service channel will work all the time. In general, the ratio of the mean of the service time distribution to the mean of the interarrival time distribution is a measure of the utilization or workload of the service channel. It should be noted that even with a low "workload" the queue can have more than one customer in it being served. The mathematics of queueing theory allows the calculation of the probabilities of such events.

In order to apply queueing theory to scanning behaviour one must have distributions of service time and interarrival time. The models of monitoring which follow provide such distributions.

An information source will from time to time "demand" attention from an observer, and, if possible, the observer will pay attention and look at that information source. On the basis of either theoretical considerations or of actual data from the real world, one can construct probability distributions of intervals between the attentional demands made by an instrument. Although it is conceivable that an instrument (or its internal representation) would demand attention on a periodic basis, in general instruments will demand attention at intervals which vary according to the characteristics of the signal driving the instrument and to the nature of the observer's task using that instrument.

Similarly the distribution of durations of fixation on an instrument can be calculated or estimated on the basis of the characteristics of the signal driving the instrument and the characteristics of the task of the observer.

It is clear that if there were some rule governing the distributions of interval and duration, so that there would be no overlap of demand, there would then be no degradation of performance on any element of a task embedded in a whole task compared with performance on that element if it were alone. Naturally such an exclusion rule would have also to include the stipulation that there would be no interferences in memory between task elements. An alternate way of achieving the non-overlapping condition would be for the observer to have the strategy of attending to the instrument about which there was the greatest uncertainty. In this case, the queue would always be of length 1.0 and the channel would always be busy. The task would be self-paced. In general, there is overlap of demand, and in general, there is interference : the human observer will be busy observing A when B demands attention. As either of two conditions arises : an increase in the frequency with which demands are made by one or more instruments, or an increase in the durations of observation times, resulting perhaps from an increase in the complexity of the signal to be monitored, the amount of interference will increase. Only if there were an increase in the threshold level of uncertainty at which observing is demanded would there be an avoidance of an increase in the amount of interference between instruments. Such an increase would be tantamount of course, to a change in the rules of the task, a relaxation of the level of precision of control or reduced awareness of system state. Under overload conditions, as commonly understood, this might be an appropriate strategy for the observer.

In summary, the instruments are the customers; the assumed single channel of visual attention is the service channel. The instruments arrive for the attention (service) and when the uncertainty about what is displayed on that instrument is reduced to zero or some sufficiently low value, the instrument leaves the service channel (is no longer fixated or no longer attended to) and something else takes its place.

If the queueing model is to be useful, there must be means of calculating the distributions of durations and frequencies of observation of the various instruments which the observer is trying to monitor. The models that do this differ in the assumptions made as to the goals of the human operator. With different assumptions, different quantitative predictions will be made of the duration and of the frequency of observation. The models generate different numbers which must be entered into the queueing model if we are to understand the consequences of each. No one model will fit all cases. There will be differences among individuals, between systems, and between tasks which will, in all probability, favour one model or another. Even within a given task, on a specified system, the human observer may shift from one model to another.

3. The Behavioural Strategies

The various models depend upon diverse definitions of the goals of the monitor. For every goal, there will, in general, be a quite different model. The goals vary in their degree of "realism", the degree to which they seem rational. The most elementary goal (and perhaps the most unreal) is that of *signal reconstruction.* The monitor is assumed to be trying to put the signal together on the basis of the samples taken on successive looks. A strategy equally unreal assumes that the choice of the next instrument to be fixated is based on a draw from an urn filled with markers. The markers are in proportion to the information generation rates of the various instruments. Thus an instrument that varied rapidly and had to be read to a high degree of precision would have more markers than some instrument that also varied rapidly but required no great precision of readout.

More rational and realistic models arise on more rational and realistic assumptions about the monitor's behaviour. One such assumption is that the monitor attends to the instrument when the *probability that the displayed value of the variable has exceeded some limit is at a maximum.* Underlying this goal, of course, is the assumption that the monitor does something only when the variable being monitored exceeds a meaningful limit. A similar assumption is, for the same task, that the monitor attends to a signal when the *probability that the displayed variable has exceeded some limit is itself greater than a probability threshold.* This will in general produce more frequent samples than the former assumption.

Yet another strategy combines the reconstruction goal with a value-dependent process. The observer is assumed to be *engaged in reconstruction, but the precision of the reconstruction is a function of how close to the limit value the previously read-out value was.*

Each behavioural strategy generates a different mathematical analysis. In all cases the assumptions about the signals will be the same : the signals displayed are random, band-limited time functions with Gaussian amplitude density distributions. The signals which drive the instruments in an array are assumed to be statistically independent and uncorrelated with one another. There are always assumed to be three or more instruments in the array.

An instrument is, of course, merely a kind of information source. Its principal virtue in this discussion is that it is a familiar and convenient way of presenting information for people to gather. The pointer movements faithfully reflect the statistical characteristics of the signal in analog fashion. In the "conditional" models, there is some limit value of the variable to be monitored beyond which the observer does not want the displayed value to go without its being observed. The analog is that of the human pilot who wishes to exert control over a variable only if it departs from the desired value by some significant amount; or perhaps that of the automobile driver who makes no corrective motion of the steering wheel unless the vehicle approaches the boundaries of the lane or the road.

4. *Models of Visual Sampling Behaviour*

A. *A Periodic Sampler (PSM)*

Consider the case of a single instrument used by an ideal observer with the goal of signal reconstruction on the basis of the samples taken. The analysis derives from Shannon (3).

The pointer of instrument i will generate a sequence of positions in time, $f_i(t)$ From $f_i(t)$ we can compute a power density spectrum $phi_i(w)$. Assume that $phi_i(w)$, has a cutoff frequency at w_i . The minimum sampling rate for periodically taken samples of $f_i(t)$ is $2w_i$, if $f_i(t)$ is to be specified from the sampless.

We can also compute the information generation rate of instrument i . Assume the observer has a permissible root mean square error of readout, E_i , and the signal has some root mean square amplitude, A_i . For $f_i(t)$ with a cutoff frequency of w_i , an rms amplitude of A_i , and an rms permissible error, E_i , the information generation rate is :

$$\dot{H}=w_i \log_2 \frac{\bar{A}_i^2}{\bar{E}_i^2} \quad \text{bits per second} \tag{1}$$

Frequencies and Durations of Fixation

Our ideal observer, samples at a rate which permits reconstrution of the signal from the samples. Therefore, H.O. must sample with frequency FF_i , which is at least equal to $2w_i$ If FF_i is exactly $2w_i$, the average amount of information which H.O. must assimilate at each sampling, $\overline{H}_i$, is :

$$\overline{H}_i = \log_2 \frac{A_i}{E_i} \qquad \text{bits per sample} \tag{2}$$

If we assume that the observer has a fixed and limited input capacity the time required for a sample, D_i , can be calculated :

$$\overline{D}_i = K \log_2 \frac{A_i}{E_i} + C \qquad \text{seconds per sample,} \tag{3}$$

where K has the dimension time per bit, and C (with the dimension second per sample) is a constant which accounts for movement time and minimum fixation time.

These are intuitively satisfying results : the fixation time is a function of the ratio of the signal amplitude to the size of the permissible error of readout; the frequency of fixation is a function of how rapidly the signal varies.

For the conditions specified, the "time demand" placed on the ideal observer is the product of fixation frequency and fixation duration, T_i :

$$T_i = \overline{FF}_i \times \overline{D}_i = 2Kw_i \log_2 \frac{A_i}{E_i} + 2w_i C \qquad \text{seconds per second} \tag{4}$$

$$= K\dot{H} + 2w_i C \tag{5}$$

T_i , the time required to be spent of instrument i, is, as it should be, a direct function of the information generation rate of the instrument i . This is obvious , of course. What is important is that the analysis allows us to predict what the frequencies and durations of fixation on instrument i will be, and to relate these predictions to data obtained in the field and in the laboratory. If fixations are taken at intervals shorter than those predicted, the amount of information to be assimilated at each fixation will be less than $\log_2 A_i / E_i$ and the duration of fixation reduced. Because of the additive constant, C , the time to be allocated to instrument i can be minimized by making $FF_i = 2w_i$.

For a complex array of m instruments, we can calculate the monitoring load placed on the observer by summing the individual times of the m instruments. The assumption of linearity is justified because of the prior assumption, axiomatic in strength, that the observer can look at only one instrument at a time.

For each instrument we calculate or measure w_i and A_i/E_i. From these we find the individual loads and sum over instruments. The sum will be the minimum utilization time per unit time for the m instruments in the array.

$$\text{Min } T_m = 2 \sum_{i=1}^{m} w_i \left(\log_2 \frac{A_i}{E_i} + C\right) \tag{6}$$

This result can be used in the design of instrument panels. For example, if a decision must be made about the addition of an instrument, we might proceed as follows : let T be the unit time; then if $T >$ Min T_m, one can try to add instrument j to the array. The values of w_j and A_j can be determined or estimated from known parameters of the system to be monitored or controlled; E_j can be determined or estimated from system performance requirements. Then the decision to add or not to add can be made rationally : if $T_j +$ Min $T_m \leq T$, add. Of course if the observer is not ideal, and does not sample so as to achieve the minimum number of eye movements, the time taken by the array will be greater than Min T_m and the calculation would have to take the inefficiency into account.

Since the underlying assumption of periodicity is untenable, the model cannot be expected to predict exactly the durations and intervals of fixations. However, it must be expected that there would be a strong ordinal relationship between bandwidth and sampling frequency since only a rescaling of time is involved in going from one bandwidth to another. The predictions should all be more "risky" than the actual behaviour of real monitors since the model assumes sampling at the Nyquist* moment for each signal. As will be seen in the elaboration of the other models, this would ordinarily be the *last* moment at which a sample might be taken. If sampling is delayed beyond T_N, the uncertainty in the monitor about the nature of the external signal has already reached a maximum, and the risk of undetected critical events has become large.

The duration of observation should be well predicted by the model on the assumption of a fixed channel capacity. If a fixation is made sooner than the model would have done, the uncertainty accumulated since the preceding fixation would be less, and the observation time required correspondingly reduced.

* The Nyquist interval (T_N) is that time displacement at which the autocorrelation function of the signal time function goes to zero for the first time. Thus a sample taken T_N after an earlier sample will be statistically independent of that earlier sample.

Fixation Sequences and Transition Probabilities : As a consequence of the sampling performed by the observer on each of the many instruments of an array, there will be eye movements which redirect the gaze from one instrument to another. In the classic work of Fitts et al (1) these "connections" between instruments on the instrument panel were called "Links". The proportion of the total number of links that connected a pair of instruments was called the "Link Value" of that pair of instruments. Thus, there was a bi-directional Link Value for every possible pair. The usage is in accord with task analysis in an extended work area. The Link Values are, of course, estimates of the transition probabilities of the instrument matrix, which have been derived from measurement in flight, in diverse flight situations. The early studies were completely empirical and, because of their atheoretical basis, treated Link Values as independent of the frequencies of fixation, It must be seen at once, however, that if two instruments have large fixation frequencies, there will, per force, be a large Link Value connecting them. This, of course, leads the periodic model into an unsolvable dilemma. Only in extremely rare circumstances would it be possible for strictly periodic sampling to take place on a multitude of instruments in an operational task. The periods would have to be commensurable and of such size as to permit a repeated sequence to occur, taking into account the duration of fixation on each instrument. Of course, it is the case that visual sampling is aperiodic and situation specific. Thus the periodic sampling model cannot be a representation of reality. It does, however, provide a boundary, based on the signal statistics, within which sampling should occur no matter what the nature of the actual process. This boundary holds, of course, only if the assumption is correct that the observer samples only the value of the signal at each sampling moment and does not observe its derivatives. In this latter case, a different model and a different limit will apply.

In summary, the PSM assumes that the observer observes the value of the variable at each sampling moment, that the task is that of signal reconstruction, that the value read is of no consequence, and that the observer has perfect memory. Given these, the model predicts that the period of sampling of an instrument will be based on the bandwidth of the time function presented on it, and the duration of sampling based on the logarithm of the ratio of the RMS of the signal to the RMS of the permissible error. It does not yield any estimates of variance of either period or duration since it assumes periodicity. It cannot deal with the question of Link Values. It does provide boundary estimates of workloads and of Link Values.

B. A Random Constrained Sampler (RCM)

Again we have an array of instruments. Each is driven by a bandlimited signal with a Gaussian amplitude density distribution of mean, μ , and standard deviation, δ. The pointer on each instrument moves in accord with the signal and the task of the observer is to look at each instrument according to some rule derived from the signal characteristics.

Duration of Fixation : We assume that the observer will sample the instruments randomly. The probability that an instrument, i , will be fixated is the ratio of the bandwidth of the signal presented on that instrument to the sum of the bandwidths of all the signals. The duration of fixation on each instrument is presumed to be, as it was for the periodic sampler, proportional to signal amplitude and permissible error :

$$\overline{D} = K \log_2 \frac{A_i}{E_i} + C \qquad \text{seconds} \tag{7}$$

where K is an observer characteristic with dimensions seconds per bit and C is a constant to account for eye movement time and minimum fixation time.

It must be recalled that the assumption is made in all the models that the monitor is *reading* the actual value of the presented variable on each fixation. Naturally, if the monitor were to be engaged in check-reading, the duration would almost certaitly be less.

The observer chooses the next point of regard by selecting a marker from an urn. The urn contains markers representing each of the instruments in proportion to its bandwidth ratio.

The expected value of the interval between fixations on instrument i will be equal to the expected number of fixations on instruments other than i , times the duration of fixation on those instruments, $\overline{D}_j$ plus the duration of fixation on i , $\overline{D}_i$:

$$E(\mathrm{int}) = \overline{D}_i + \overline{D}_j \left(\frac{1-p_i}{p_i}\right) = \overline{D}_i - \overline{D}_j + \frac{\overline{D}_j}{p_i} \tag{8}$$

where $\overline{D}_i$ is the mean duration of fixation on instrument i , and $\overline{D}_j$ is the mean duration of fixation on all other instruments. Since we assume that the duration of fixation will be equal for all instruments, because of the fact that the required precision of reading is the same for all, the expected value of the interval reduces to :

$$E\ (\text{interval}) = \frac{\overline{D}}{p_{i}} \qquad \text{where } \overline{D} \text{ is the overall mean of fixation} \tag{9}$$

Given the simplifying assumption of equality of mean fixation duration on all the instruments the variance of the distribution of intervals can be shown to be :

$$\text{Var (interval)} \qquad \frac{\sigma^2_{\overline{D}}}{p_{i}} + \frac{1-p_{i}}{p_{i}^2}\ \overline{D}^2 \tag{10}$$

In accord with the distribution so defined, the instrument monitor, H.O., chooses a marker i and casts it back into the urn. H.O. then looks at instrument i for D seconds and chooses another marker.

Fixation Sequences and Transition Probabilities : This model allows the calculation of transition probabilities or Link Values .

The probability that the eye will fixate on instrument i is p_{i} , and on instrument j is p_j The probability of a transition from i to j is $p_{i}p_j$. This is also the probability of a transition from instrument j to instrument i . The Link Value is therefore :

$$P_{\overline{ij}} = 2p_{i}p_j = 2\,\frac{w_{i}w_j}{\Sigma w} \tag{11}$$

It is clear that if p_{i} and p_j are large, many transitions will perforce be made between i and j . On the other hand, as the probabilities of sampling the various instruments approach one another, the freedom of path of the observer through the array increases and becomes complete when the probabilities are equal. Thus as the constraint of relative probabilities of fixation diminish, there is greater opportunity for "logical" patterns of scanning to occur By logical is meant in accord with some scheme of information uptake related to the nature of the information presented on each instrument. For example, in an aircraft, one might expect that a pilot would shift from "altitude" to "rate-of-climb" to "engine output". In practice, such sequences do not occur importantly more often than is predicted and data collected in flight adhere remarkably well to the model (4).

Repeated Observation Effects : If the observer is fixated on instrument i , there is a probability p_{i} that the next fixation will also be on instrument i , if the model is strictly interpreted. This very much affects the empirical data which will be obtained from measurement of a subject engaged in an observing task. Because a "shift from i to i "

will appear merely to be a longer than usual fixation on i , the measured frequency of fixation will fall short of the predicted value by $p_i \times FF_i$ fixations per unit time. As a consequence, the observed frequency of fixation of instrument i , FF_{oi} must be corrected (if it is to be compared with the prediction of the model) :

$$FF_{oi} = FF_i \ (1-p_i) = 2w_i \ (1-p_i) \quad \text{if} \quad FF_i = 2w_i \qquad (12)\ *$$

The numerical deficiency will be the greater for the instruments with the greater bandwidths (and higher frequencies of fixation) and as a result, the relative fixation frequencies for the instruments with the lesser bandwidths will be increased.

Concommitantly, there will be an increase in the observed duration of fixation, D_{oi} , because some of the fixations (according to the model) will be of double length or triple or greater length and so on. The observed duration must be corrected (if it is to be compared with the predictions of the model) :

$$\bar{D}_{oi} = \bar{D}_i \ \frac{1}{1-p_i} = \frac{w_i}{\Sigma w} \ K \log_2 \frac{A_i}{\Sigma i} + \frac{w_i}{\Sigma w} \ C \qquad (13)$$

The model predicts, therefore, that there should be a positive correlation between the bandwidth of information presented on an instrument and the mean duration of fixation on that instrument even with identical requirements for precision of reading.

The observable Link Values, $p_{o\overline{ij}}$, will be less than $2p_i . P_j$: $p_{o\overline{ij}} = \dfrac{2p_i . p_j}{1 \sum_1^m p^2}$ (14)

In summary, the RCM assumes that the observer chooses from the array of instruments one with a probability which is a function of the bandwidth of that instrument compared with the bandwidths of all the other instruments. The frequency of sampling is merely the probability of each instrument divided by the duration of observation averaged over all instruments. The model offers a variance estimate of the periods between observation of each instrument.

* There is no logical barrier to an infinite number of "self-transitions" with ever diminishing probability.
The sum of the infinite number of terms : $\sum_\infty p^n = \dfrac{1}{1-p}$
hence the multipliers l-p and l/l-p for the observed data.

The model predicts that the frequencies, durations and Link Values found in the experimental data will be altered by virtue of the selftransitions. The larger the bandwidth of an instrument, the greater the probability of its being fixated and the greater the effect on the observable data of the self-transitions. Durations of observation will be increased, frequencies of observation decreased. Link Values will decrease for the greater bandwidths and increase for the lesser bandwidths.

However, the queue always has but a single awaiting service or being served. No new signal may join the queue until the previous signal has been serviced.

C. Conditional Sampling Models : General.

Conditional sampling is a process quite different from the periodic sampling analyzed by Shannon (3) and used in the first model. Here the timing of the next sample to be taken is a rational function of what was read on the preceding sample. An instrument monitor engaged in "check reading" takes no action with respect to a value of the signal unless it exceeds some boundary condition of acceptable system behaviour, except, possibly, to compute the next moment of sampling.

A homely example which serves to illustrate the process is that of baby watching. Imagine yourself to be trying to read this monograph whilst seated on the lawn near a swimming pool. An infant is crawling on the grass generating a "random crawl". You wish to intervene when the infant is likely to fall into the pool and you wish to get as much reading done as possible, as well. A sensible strategy would be to calculate a next time to look at the infant based on what you had observed on the last look. If the infant had been close to the pool's edge, you would look much sooner than if it had been far away. Other things being equal, you would look sooner if it had been approaching the edge of the pool than if it had been receding from it. Lastly, in general and other things being equal, you would look sooner if it were an active infant than if it were a lethargic one. Thus the determinants of your observing behaviour would be : the amount by which the value observed fell short of the limit; the derivative of the observed variable; and the mean absolute velocity of the variable (which will be a direct function of the bandwidth of the signal formed by the positions of the infant in time).

The instrument monitor is solving the same problem; and the same variables will influence his behaviour. The difference is that he has a large number of "infants" to watch and he must distribute his visual attention amongst them in a way which ensure that the possibility of a "drowning" is minimized. What should his strategy be ?

A number of possible strategies will be assumed and the consequences of each examined. Obviously there are some over-riding matters which govern behaviour no matter what the strategy. One of these is the question of the required accuracy of read-out of the value of the signal presented on the instrument. The same rule should hold as in the case of the simple periodic sampler : if the ratio of signal power to permissible error is small, the durations of samples will be small. If the ratio is large, then the durations will be large. It might also be rationally assumed that the required accuracy will vary as a function of the distance of the value presented from the limit value for that instrument. Then as the variable approaches the limit, the duration of observation will increase. An example from automobile driving illustrates the relationship. As the gasoline gauge progresses from "full" to "empty", the gauge is read more and more precisely since the consequence of and given size of error increases In general, also, it is immediately clear that as the displayed value approaches the limit value, the demand of the instrument increases since the interval becomes shorter as the duration becomes longer. When the value is at or about the limit, the observer will spend all his time observing it unless some other emergency strategy takes over which involves corrective action. In most cases when monitoring discloses extreme values, control behaviour takes over until the variables are brought back within limits.

D. Conditional Sampling Model I - Sample when probability of exceeding limit is maximum. (CSM I)

The assumptions underlying the model are :

1. The displayed signal is a random, bandlimited, Gaussian signal with mean 0.0, and variance σ^2 .

2. There is some limit L beyond which the monitor does not want the signal to go without its being observed.

3. The cost of taking a sample is small compared to the cost of exceeding the limit.

With these assumptions, the signal will have an autocorrelation function

$$p(t) = \frac{\sin 2\pi W_c t}{2\pi W_c t} \tag{15}$$

Where W_c is the cut-off frequency in radians per second and t is the ratio τ/T_N, τ being the shift between the signal at time 0 and time T_N

The autocorrelation is 1.0 at $t = 0.0$ and 0.0 at $t = T_N$.

All computation will be in terms of this non-dimensional time : percent of T_N.

As t increases from 0.0 to T_N, the expected value of the variable changes from x(0) (the previously read value) :

$$\mu(t) = E\ (x(t)) = \rho(t) \cdot x\ (0) \tag{16}$$

and the standard deviation of the variable around $x(t)$ increases with decreasing $\rho(t)$:

$$\sigma_x\ (t) = \sigma_x\ (1-\rho^2\ (t))^{\frac{1}{2}} \tag{17}$$

Then the probability that $x(t)$ will exceed some limit, L, is greatest when the distance from (t) to L in units of $\sigma_x(t)$ is least. The z score of L with respect to the expected value of x at time t will be :

$$z_L\ (t) = \frac{L-(0)\ \rho(t)}{\sigma_x(1-\rho^2(t)^{\frac{1}{2}}} \tag{18}$$

and $z_L(t)$ will be a minimum and the probability a maximum when the derivative of $z_L(t)$ with respect to σ is zero. Thus :

$$\frac{\partial z_L(t)}{\partial \rho(t)} = \frac{L\rho(t)\ -\ x\ (0)}{\sigma_x(1-\rho^2(t))^{3/2}} \tag{19}$$

Then $\rho(t)$ which yields $z_L(t)$ a minimum, and the probability that $x(t)$ will exceed L a maximum, will be :

$$\rho(t) = \frac{x(0)}{L} \tag{20}$$

As x(0) approaches L, $\rho(t)$ approaches 1.0 and t approaches 0.0. As x(0) approaches the mean, 0.0 , the sampling interval, t , goes to the Nyquist interval.

Naturally, if x(0) is $\geq$ L, the next observation will be taken without delay, i.e., observation will be continuous.

Conditional Sampling Model 2 - Sample when probability of exceeding limit exceeds a probability threshold. (CSM-2)

We may, instead of seeking a maximum probability of exceeding the limit L, choose to sample when the probability is itself equal to or greater than some arbitrary probability. Since the amplitude density distribution of the signal is assumed Gaussian, we can characterize this probability level by the z-score associated with it. A rational choice of probability threshold is that which holds at T_N , when the expected value of $x(t)$ is μ and the variance is σ^2_x At that moment the criterion will merely be z_L . For values of t between 0.0. and T_N , the expected value of the variable will be $\mu(t)=x(0)\rho(t)$ and the variance will be $\sigma^2_x(1-\rho^2(t))$ The z-score of L, $z_L(t)$ will be a function of $\rho(t)$:

$$z_L(t) = \frac{L-x(0)\ \rho(t)}{\sigma_x(1-\rho^2\ (t))^{\frac{1}{2}}} \tag{21}$$

Since we wish to sample when the probability of exceeding L is equal to the probability at T_N , we can set $z_L(t)$ equal to z_L :

$$\frac{L-x(0)\ \rho(t)}{\sigma_x(1-\rho^2(t))^{\frac{1}{2}}} = \frac{L}{\sigma_x} \tag{22}$$

and solve for the value of the correlation coefficient, $\rho(t)$, in terms of x(0) and L :

$$\rho(t) = \frac{2Lx(0)}{x^2(0) + L^2} \tag{23}$$

Once again we can calculate the fraction of the Nyquist interval corresponding to any value of $\rho(t)$. The greater the distance between x(0) and L, the greater the interval to the next sample, and v.v.

A comparison of this result with that of CSM-1 shows that for any value of x(0) the sample is taken earlier than for the maximum probability, This is as it should be, of course.

F. *Conditional Sampling Model 3 - A variable Nyquist sampler.* (CSM 3)

Given the same assumptions about signal statistics with the provision that the signal power declines linearly after some frequency, we can construct a sampler which has the goal of reconstruction of the signal but varies the interval between samples on the basis of a varying bandwidth.

The effective bandwidth will be a function of the permissible error between signal and readout. As the power of the permissible error increases, the frequency above which the signal power is less than the error power is lower. If permissible error becomes smaller, the frequency above which the signal power is less than the error power is higher. Given the assumption of linear decline in signal power we can calculate the effective cut-off frequency and, hence the sampling interval. The precise relationship cannot be found without a specification of the decline in power per octave and the relative size of the permissible error.

Durations of Samples

The duration of each fixation can be calculated in a number of ways. The variance of the signal increases with time :

$$\sigma_x^2\ (t) = \sigma_x^2\ (1-\rho^2(t)) \tag{24}$$

and the information to be taken up increases with time :

$$H(t) = \log_2\sqrt{2\pi e\sigma_x^2\ (1-\rho^2(t))} \quad \text{bits} \tag{25}$$

and fixation time will be :

$$D(t) = K\ H(t) + C \quad \text{seconds} \tag{26}$$

and duration of fixation is a variable.

Alternatively we can assume that the monitor reads each signal to a constant precision, *F*. The information to be taken up is constant in time ;

$$H = \log_2 F \quad \text{bits} \tag{27}$$

and fixation time will be constant except for internally generated variability :

$$D = K \log_2 F + C \quad \text{seconds} \tag{28}$$

Other relationships can be logically derived. Only experimental data will enable a correct choice to be made.

All the models predict both intervals and durations of sampling of each instrument in an array. Since the moment of sampling for the conditional sampling models has been scaled in non-dimensional time t/T_N, it is clear that the predictions of interval are in strict relationship to bandwidth. Ceteris paribus, as bandwidth varies, so does interval and in just proportion. The duration predictions are in general accord with the notion of a limited and constant channel capacity which yields an observation time in linear relation to the amount of information to be gained during the fixation.

G. A Signal Reconstructor with Imperfect memory - "Rational Model" (RSM)

The periodic sampler was orginally constructed as a simple and unrealistic model of human behaviour. It yielded approximations of sampling behaviour in the laboratory but failed in the analysis of data from simulated operational tasks. It is possible, in light of the later development of Fogel (5) to construct a better model resting on the same logic but with more realistic assumptions.

Let the strategy of the sampler be to reconstruct the displayed signals on the basis of samples taken at intervals . The signals are Gaussian, bandlimited time functions. Naturally, the rate of change of the signal varies as a function of time. It is the case that the amplitude density distribution of a bandlimited Gaussian function is itself a Gaussian time function of the same characteristics. The higher rates of change will be above some velocity threshold of the observer at some bandwidth of the displayed signal. The higher the bandwidth, the greater proportion of the total signal will have rates of change above the observer's threshold. If the instantaneous velocity is sampled by the observer, the next sample may be delayed for an interval twice as great as the Nyquist interval. Thus, the sampling frequency for signals whose velocity can be perceived will be half the sampling frequency for signals for which the rates of change are below threshold. Since the proportion of suprathreshold velocities will increase with increasing bandwidth, the slope for higher bandwidth signals will become progressively less, decreasing from 2.0. to 1.0 over some range. Each point on the sampling frequency versus signal bandwidth function will be on a !ine of different slope.

For realistic modelling of the human operator, we cannot assume that there is perfect memory but must allow forgetting. Thus, for the low bandwidth signals, the interval calculated on the basis of periodic sampling will be in excess of the forgetting time. The dominant factor influencing the interval for low bandwidth signals will be the time constant of memory loss. The relation of sampling frequency and bandwidth will become flat and asymptote at a horizontal line at a frequency which is the reciprocal of the forgetting time constant.

For the higher bandwidth signals the data will show a transition to a rising function along the slope 2.0 line and then with increasing signal bandwidth become concave downwards until the data intersect the slope 1.0 line. Then the data stay on that line. (Unless we postulate that the observer can sense acceleration on some occasions for the higher bandwidth signals. In which case there will be a continuing concavity downward until the line with slope 2/3 is reached).

Since the rate of change of the displayed variable changes from observation to observation, the model yields variable sampling behaviour. If there were a sharp threshold between movement seen and movement not seen, the data would fall into two heaps, one on the line of slope 2.0, and other on the line of slope I.0. Since, of course, there is no exact threshold "below which nothing", there would be a continuous variation of the interval as the amount of velocity information extracted from the signal varied from observation to observation. In particular, we would expect that the distribution of sampling intervals would be skewed positively for the lower bandwidths, symmetric for the central bandwidths, and skewed negatively for the higher bandwidths.

Since we have no way at this time of finding the velocities which had been displayed in the following experiments we cannot make a complete comparison of the predictions of the "rational" sampler and the empirical data.

5. Discussion

The various models predict different variances for the data. The Periodic Sampler predicts no variability at all, a clearly unrealistic prediction. Each of the others has a variance. In the case of the Random Constrained Sampler, the variance is the consequence of the sampling process. In the Conditional Samplers, the interval is assumed to be a determined function of the previously observed value. This, however, is itself a random variable with a variance which in turn depends on the timing of the subsequent sample.

Although the underlying distribution is assumed to be Gaussian, the conditional nature of the sampling process makes the distribution of observed values, and the consequent distribution of sampling intervals, non-Gaussian. The extreme values near the limit are sampled more often because of shortness of interval and the consequent high values of autocorrelation between previous and subsequent samples. The distribution of observed values becomes non-Gaussian and more "rectangularized" from the monitor's point of view. This may not significantly alter the subjective uncertainty about the signals.

It is possible to ignore the individual instruments and to calculate the time which will be required to monitor an array on the basis of the total information flow from that array. Since such calculations ignore the time required both to choose the next point of regard and the time of the eye movement itself, the loads will be too low. There are some circumstances in which the individual sources of information cannot be identified. Automobile driving is one such. Here it is necessary to consider the global information processing load rather than the sum of the components, except for the two major categories of uncertainty : internal and external. A complete analysis of a model of this sort is presented in Appendix B.

The models which have been presented provide measures of estimates of the loads of each of the instruments. If one assumes that these loads are orthogonal, as we have done, then the vectorial workload model presented in Appendix C can be used to estimate the total workload. An alternative would be to use the queuing model which is presented in Appendix D to obtain estimates of the utillization of the observer considered as a service channel with many customers.

The experiments which follow were composed while the model development was growing from the simplest to the more complex and more realistic. In order to test the conditional models, it would be necessary to know what the instrument said on each observation, and, to know how accurately the displayed value was read. The former is difficult although not impossible. At the time that the experiments were done, it could not have been done with the available instrumentation. The latter, to know how accurately the observer knew what was shown him, is probably impossible. Tests to dertmine which model best predicts the data can be based, therefore, only on the statistics generated and not on the succession of intervals. Yet it is the succession of intervals which is predicted and which distinguishes between the various models. The tests, therefore, are weak. More sophisticated experiments will be required to test which model is closest to reality.

Even, then, we have no assurance that different observers do not use different processes, nor that an observer at different times may not use different processes, either at different stages of learning or at different levels of arousal, motivation or what have you. Each of the models, except perhaps for the Periodic Sampler, has a basic rationality and does not seem implausible for someone somewhere.

CHAPTER III

EXPERIMENTS ON VISUAL SCANNING BEHAVIOUR

1. Introduction

The array of models described in Chapter II gives rise to a great number of experimental hypotheses which could be tested. The development of the models took place over a span of many years. The experiments which are reported herein were conceived at the time of the exposition of the periodic sampling model, or shortly thereafter. While the Six Dial Experiment was being prepared and carried out, the Conditional Sampling Models were elaborated. Still later, the Random Constrained Model was formulated to clarify its relationship to the Periodic Model, and the Conditional Sampling Models were completely reformulated to present a more easily understood and more easily calculated mathematical structure. It is possible to calculate the predictions of the Random Constrained Model, even though it was formulated long after the experiments themselves were run and analyzed, and this will be done where it is feasible.

At the time of the first modelling and experimentation (1953-1958), the technology of eye movement recording was relatively crude. Electrooculography was well established but drift of the recordable potentials at the surface of the skin was a serious problem if one wished to record the position of the eye in absolute angular measure. The cumbersome but reliable motion picture technique which had been used, and was then still being used for the in-flight recording of eye positions, seemed to be the best of the methods. One of the drawbacks of the method, which had to be accepted, was that the film had to be read by teams of persons who compared each frame of motion-picture film with standardized reference photos which calibrated the appearance of the recorded eye positions. It would have been possible, although difficult, to record simultaneously the readings of all the dials which the subjects were observing. Since the theoretical position at the time was that the monitoring behaviour was a function of the statistical characteristics of the signal rather than conditional on the actual value of the variable read on each fixation, these values were not recorded.

The experiments must therefore be considered as appropriate only to the models which do not depend on a knowledge on the part of the observer of what the instrument which was fixated actually said. Of course all of the models predict that there would be an increase in sampling frequency with an increase in bandwidth. The underlying assumption that the monitor is engaged in uncertainty reduction through the acquisition of information from the displays mandates, under a state of high loading, that more time, hence more fixations, should be devoted to the signals which present more information. In these experiments the information generation rate was a linear increasing function of the bandwidth of the signal.

The Periodic, the Random and the Conditional models all predict a linear increase in sampling frequency with increase in bandwidth. The Rate Detector with imperfect Memory predicts a curvilinear relationship between signal bandwidth and sampling frequency. Thus the results should enable us to discriminate between the two kinds of model.

If the data do not force the rejection of linearity, then an examination of the slope of the best-fitting line will permit some degree of decision as to which of the linear predictors is closest to the data.

2. *Experiment 1 : Four-Instrument Scanning*

The purpose of this experiment was to examine the sampling behaviour of subjects watching instruments presenting signals of diverse bandwidths. The experimental hypothesis was that the frequency of fixation on each of the instruments would be linearly related to the bandwidth of the signals with a slope of 2.0 as shown in Chapter II. In other words, the Periodic Sampling Model predicts the sampling frequency. Naturally, it was not expected that the sampling would be periodic in fact. However, on theoretical grounds it seemed that the average sampling frequency should conform to the predictions of the model. If samples were taken later than the Nyquist interval, the observer would not have sufficient information; if the samples were to be taken earlier, there would be a waste of time. The observer should therefore tend, if the sampling load were high enough, toward the optimum solution of the task : to sample very nearly periodically and with a mean of twice the bandwidth of each signal.

Method :

The photographic technique for measurement of fixations was selected for its lack of drift and its high degree of face validity with respect to point of fixation. The apparatus was available and trained film readers as well. The instruments were mounted on panels at a separation sufficient to prevent peripheral reading of the dials. The instruments themselves were micro-ammeters with center zero and scaled to plus and minus 50 at about 70 degrees on either side of zero. The 3 inch diameter subtended an angle of 6 degrees at the eye of the observer. The scale markings were commercial standard white on a black background. The pointer was finer than would ordinarily be recommended for ease of reading but since one object was to minimize peripheral uptake of information, this was not considered to be a drawback. The instrument array was illuminated with incandescent light well into the photopic range at about 50 foot candles. The array was arranged with an instrument at each corner of a 23 degree's (of visual angle) square. In the center of the square was a viewing hole for the camera lens. The camera was a motion-picture camera, with a 25mm telephoto lens. A chin-rest kept the subject's eyes within the field of the ciné camera. This combination permitted the recording of both eyes and a portion of the surrounding face. The camera ran at 8 frames per second as had been done with the similar apparatus in the Pilot Eye Movement Studies of Milton, Jones and Fitts (I). Tri-X film was used with an ASA rating of 400. This permitted good recording of the subject's "mask" without the need for very high levels of illumination on the subject's face.

Subjects :

The subject were five Antioch College undergraduates. They volunteered for the experiment and were paid $ 1.25 per hour for their time. A bonus was paid for detection rates (see Procedure) over 95% . All the subjects completed the experimental program.

Procedure :

Each subject was seated in front of the instrument array at a distance of 72 cm. Each 76 mm dial subtended, therefore, an angle of approximately 6 degrees. The subjects were instructed to watch the instruments and to press a hand switch, held in the hand of choice, whenever *any* of the pointers was at 40 or more on either side of the dial. They were also instructed to press the switch if they thought that any pointer *had been* at 40 or more during the time since the last time it had been looked at, or if they expected it to go "out of bounds" immediately after they looked at it and were looking away. Three minutes of camera time were obtained at the beginning and at the end of each one hour session. There was one session on each of 30 successive days.

The arrangements of signals were different for each of the four subjects. The instrument readings themselves were recorded so that the actual events of 40 or more could be seen by the data analyst and the bonus calculated. The responses occurring within I second of the exceeding of 40 were counted. False alarms were not counted but there were few.

The Theory of Signal Detection (TSD) was not in common use at the time so that it did not seem necessary to record the false alarms and to analyze in accord with TSD. Naturally, in retrospect, it seems reasonable to do such an analysis in order to gain more insight into the internal workings of the subjects' decision processes.

Data Reduction and Analysis :

The films were developed and read by two independent film readers. Each frame was compared with reference photos of that subject looking at each of the four instruments and its probable fixation point determined. The results of the two readers were compared and differences were resolved by discussion and finally, if needed, decision by the experimenter. There were few differences as is characteristic of the method. The number of frames from one fixation to the next fixation on the same instrument was recorded for each of the instruments, as the inter-fixation interval. These data were calculated for each instrument and for each subject. The intervals were converted into frequencies and plotted as a function of signal bandwidth. Best fitting straight lines were calculated for each of the subjects and for the group.

Results :

The outcome of the experiment is described in Appendix A which follows immediately. In general, there was strong support for the experimental hypothesis. The sampling frequencies increased strongly with increasing signal bandwidth. The subjects' data were generally above the theoretical relationship. The slopes varied. Some were greather than, some less than 2.0 .

Discussion :

The experiment was, in light of modern techniques of electrooculography, and computer data gathering and analysis, *primitive,* and *naive.* The temporal resolution of the interval measurement was woefully coarse. The statistical analysis was barely adequate to the quality of the data. More could have been done although at the time the experiment was monumental in its demands on equipment and personnel. It did demonstrate that there was a strong positive, monotonic relationship between frequency of fixation on an instrument and the bandwidth of the signal presented on that instrument. As a useful engineering technique for the estimation of workload, and for instrument panel design, it was satisfactory even though the model was manifestly unreal. It even offered a method of evaluating the quality of an instrument or an instrument arrangement in vivo rather than in the artificial isolation of the laboratory experiment on one dial at a time. Thus if a dial were looked at much more often than the model suggested, it might be imagined that there was something wrong with the instrument or with the arrangement in which it was embedded. Also, if it were looked at much longer than the model suggested, there was probably something wrong with the design of the instrument face. Or, of course, there might in either case be something wrong with the training or intent of the human observer. But from the theoretical point of view, the model could not serve as a basis for a psychological theory of scanning or monitoring behaviour. That had to wait for the development of the Conditional Sampling Models some years later.

Part of the analysis, the Link Values calculations, did not really belong to the Periodic Sampling Model. Again, the anomaly was recognized but ignored at the time of the Four Dial Experiment. The Link Value Model was useful even if it was in direct contradiction (it predicted on the basis of selection of the next point of regard from an urn) to the hypothesis which underlay the Periodic Model. In general, it was evaluated on the basis of its utility rather than on its consistency. The general approach was an engineering one. The full text of the original publication follows immediately, as Appendix A.

APPENDIX A

The Human Operator as a Monitor and Controller of Multidegree of Freedom Systems

ABSTRACT

Although most research and theory building on human operator performance has considered the operator to be a continuous single-channel controller, straightforward examination of real situations and real behaviour shows him to be a sampled-data, commutated single-channel controller. This arises from the fact that the human operator must distribute his attention sequentially over many information sources. Both monitoring and controlling behaviour are subject to the constraint that the operator's eyes can fixate on only one place at a time.

Whether for an analysis of the problem of control or monitoring, it is necessary to have an understanding of the sampling process exhibited by the operator. This report sets forth a model which attempts to predict the relation between the kind and rate of information displayed on any display and the frequency and duration of samples made of that display. The approach utilizes the notion that it is possible to quantify the attentional demand or work load placed on the monitor or controller by each source of information in a complex man-machine system. The attentional demand can be calculated on the basis of the bandwidth and required precision of readout of the signal presented by an information source. It can be measured by the frequency and duration of fixations on an information source. The results of theoretical calculations compare favorably with experimental results.

1. Introduction

Virtually all of the research on and analysis of the human operator over the last twenty years has dealt with single degree of freedom systems. The human operator model treated twenty years ago in a discussion of optimum-aided tracking systems is a single channel system (1). Similarly, the various succeeding approaches : Birmingham and Taylor (2), Elkind (3), (4) Mayne (5), McRuer and Krendel (6), and North (7) have all implicitly assumed continuity of input and output and, almost as a corallary, single channelness of the operator. Bekey (8) dealt with the operator as a sampled-data system, but was concerned with a hypothetical internal sampling rather than the observable external sampling. There have also been a few empirical studies of sampled-data tracking systems in which the human operator is placed. Although good data have been produced, no broad theoretical structure has been formed. Thus, Battig et al (9), Bennet (10), and Senders (11), (12) have dealt variously with intermittent displays and controls.

These studies have not considered the fact that in most tasks the human operator must distribute his attention over many displays. The distribution of attention over related display-control pairs makes of the human operator a single-channel device which is commutated in some aperiodic sequence over a number of (perhaps interrelated) closed loops. In homely terms, a pilot looks at many instruments, one at a time, and takes corrective action as needed. The pilot eye movement studies of Milton, et al. (13), demonstrate this obvious fact and quantify the distributions. Regrettably, in those studies, no record was made of the instrument readings over the experimental period. As a result, one could only speculate as to why an instrument was looked at with the obtained distribution of frequencies and durations.

2. *Theoretical Discussion*

A. *Frequencies and Durations of Sampling*

Some general (and simple) theoretical notions about the sampling behaviour of human monitors are presented here (14). It is impossible to estimate the information presented by a continiously varying instrument if consideration is given only to the instrument itself, apart from its use. It is still more difficult to estimate the total information flow from a display consisting of a multiplicity of instruments differing from one another in a variety of ways. Let us consider first the case of the single instrument (among many) as it is used by an ideal observer.

1) The Single Instrument : An instrument, i , will generate (under given system conditions) a sequence of pointer positions in time, $f_i(t)$. From $f_i(t)$ we can compute power density spectrum $\phi_i(\omega)$. Assume that $\phi_i(\omega)$ has a maximum frequency of (cutoff frequency) of W_i . The minimum sampling rate for periodically taken samples of the function $f_i(t)$ will be $2W_i$, if $f_i(t)$ is to be specifiable from the samples. We can also calculate the rate at which the instrument is generating information, if we specify a permissible rms error of readout by the observer, and the rms amplitude of the signal (15). For $f_i(t)$, with a cutoff frequency of W_i , an rms amplitude of A_i , and a permissible rms error of E_i , the information generation rate is

$$\dot{H} = W_i \log_2 \frac{A_i^2}{E_i^2} \text{ bits/sec.} \tag{1}$$

Our ideal observer samples at a rate which permits the reconstruction of the signal from the samples. Therefore, he must sample with a fixation frequency $\overline{FF}_i$, which is at least equal to $2W_i$. If FF_i is exactly equal to $2W_i$, then the average amount of information which he must assimilate at each sampling, $\overline{H}_i$, is

$$\overline{H}_i = \log_2 \frac{A_i}{E_i} \quad \text{bits.} \tag{2}$$

Reaction time has been shown by Hick (16) and Hyman (17) to increase with increasing stimulus information. For some stimulus conditions, the relationship has been shown to be linear. If we assume that our ideal observer has a fixed input channel capacity, then the duration of each fixation, D_i, should also be linearly related to the amount of information to be taken in at each observation. Therefore, we can calculate to be $\overline{D}_i$

$$\overline{D}_i = K \log_2 \frac{A_i}{E_i} + C \quad \text{sec.} \tag{3}$$

where K has the dimensions of time per bit, and C (with the dimensions : time per fixation) is a constant to acount for movement time and minimum fixation time. This is an intuitively satisfying result : A_i is related to the possible range of values which the instrument could present, and E_i is a measure of the accuracy to which the instrument must be read. For the conditions specified, the attentional demand or work load placed on our observer by instrument i is clearly the product T_i of the fixation frequency $\overline{FF}_i$ and fixation duration $\overline{D}_i$.

$$T_i = \overline{FF}_X \times \overline{D}_i = 2KW_i \log_2 \frac{A_i}{E_i} + 2W_iC \quad \text{sec./sec.} \tag{4}$$

T_i, the proportion of total time spent on instrument i, is , as it should be, related to the information generation rate of the instrument $\overline{H}_i$.

If the fixation frequency is greater than $2W_i$, the samples will be correlated and the amount of information to be taken in at each sample will be less than $\log_2 A_i/E_i$. Since $\overline{H}$ is a property of the signal and not of the sampling process, it will be constant as the fixation frequency increases. Thus

$$\overline{H}_i = \frac{2W_i}{FF_i} \times \log_2 \frac{A_i}{E_i} \quad \text{bits.} \tag{5}$$

and

$$\overline{D}_i = \frac{2KW_i}{FF_i} \times \log_2 \frac{A_i}{E_i} + C \quad \text{sec.} \tag{6}$$

Because of the additive constant C , the percentage total time spent on an instrument is minimized by making $FF_i = 2W_i$, as in (4).

2) Multiple Instrument Displays : For a complex of m instruments, we can calculate the total work load placed on the ideal observer by summing the individual work loads of the m instruments. For each instrument, we calculate or measure W_i , and A_i / E_i. From these we calculate the product $\overline{FF}_i \times \overline{D}_i$ and sum across instruments. The sum would be the minimum utilization time per unit time for the m instruments,

$$\text{Min } T_m = 2 \sum_{i=1}^{i=m} W_i \left(K \log_2 \frac{A_i}{E_i} + C\right) \tag{7}$$

This result can be used in the design of instrument panels. For example, if a decision must be made about the addition of an instrument, we might proceed as follows : let T be the unit time; then, if $T > \text{Min } T_m$, one can try to add instrument j to the set of instruments. W_i and A_i can be determined of estimated from known parameters or the system to be monitored or controlled; E_i can be determined or estimated from the system requirements. Therefore, the decision to add or not to add can be made rationally : if $T_i + \text{Min } T_m \leq T$, add.

B. Fixation Sequences

As a consequence of the sampling performed by the observer on the various instruments of a set, transitions will be made from one instrument to another and frequency distributions of such transitions will be generated.

Transition Probabilities : We can examine the consequences of the assumption that the sequence of transitions is a random series constrained only by the relative frequencies of fixation of the instruments involved in any transition. We assume that a transition starting from instrument i may end on any instrument, including instrument i , in accord with the probabilities of fixation on each instrument. Over a sufficiently long time interval, the relative number of fixations on each instrument will be an estimate of the probability of fixation on that instrument, and this in turn reduces to the equating of the relative frequency of fixation to the probability of fixation . Thus,

$$P_i = \frac{\dfrac{T \times FF_i}{N}}{T \sum\limits_{i=1} FF_i} = \frac{\dfrac{FF_i}{N}}{\sum\limits_{i=1} FF_i} \tag{8}$$

The probability of a transition between instrument a and instrument b is $P_a P_b$; the probability of transitions in both directions, $P_{\overline{ab}}$, is

$$P_{\overline{ab}} = 2P_a P_b \tag{9}$$

It is clear that if P_a and P_b are large, many transitions will perforce be made between them. However, it is also obvious that, as the probabilities of the various instruments approach one another, the freedom of path through the set of instruments increases and is maximal when all are equal. Thus, as the restraints of relative frequency diminish, there is greater opportunity for logical patterns of scanning to occur. We expect, however, that much of what has been observed about transition probabilities can be calculated on the basis of the sampling frequencies.

C. *Measureable Data*

If the observer is looking at instrument *a*, there is a probability P_a that the next observation will also be on instrument *a*. This fact very much affects the empirical data which will be obtained from measurements of a multi-instrument task. In the first place, the measured frequency of observation will fall short of that predicted by $P_a \cdot FF_a$ samples per second. The observable frequency of observation of instrument a, FF_{0a} must be corrected :

$$FF_{0a} = FF_a\ (1 - P_a) - 2W_a(1 - P_a)\ \text{if}\ \ FF_a = 2W_a. \tag{10}$$

The numerical deficiency will be proportionally larger for the instruments with the greater bandwidths (and higher frequencies of fixation) and as a result, the relative frequencies for the instruments with the lesser bandwidths will be increased.

In the second place, the pair of observations of *a* constitutes an unobservable transition from *a* to *a* which occurs with probability P_a^2. Therefore the observable probability of transition from *a* to *b*, $P_{\overline{a0b}}$ must be corrected .

$$P_{0ab} = 2P_a P_b - \sum_{i=1}^{N} (P^2_{\ i}) \tag{11}$$

Therefore, the observable transition probabilities will be larger than those calculated on the basis of (9).

By the same process, the distribution of observable durations of fixation will be skewed toward larger values, and the observable mean duration of fixation $\overline{D}_{0a}$ must be corrected :

$$\overline{D}_{0a} = \left(\frac{1}{1 - P_a}\right) \left(K \log_2 \frac{A_a}{E_a} + C\right). \tag{12}$$

III. Experimental Data.

Data are available from a preliminary experiment which afford a partial test of the model presented in Section II (18).

A. The Preliminary Experiment

The task of the subjects was that of observing the bank of dials and making a trigger-pressing response when any of the pointers exceeded a predetermined set point on the dial. The same response and the same trigger were used for all dials; no selection of response was required for different dials. A record was taken of responses as well as the state of the dials during the sampling period.

1) Experimental Conditions : The apparatus consisted of a 4-channel random function generator adjusted to provide four random functions with bandwidths of 1/2, 1, 2, and 4 radians per second. Each channel was connected to a separate microammeter. The subject was positioned facing a panel at a distance of 28 inches. The four meters were at the corners of a square 12 inches on the side centered in the visual field. At the center of the square was a motion picture camera focused on the eyes of the subject and recording at eight frames per second during that portion of each trial when data were taken. A total of five subjects was used and each subject monitored the bank of four dials for one hour per day for 30 days. A three-minute sample of every hour was recorded in its entirety, and the results from these samples constitute the data of the experiment.

2) Scoring : A response was scored as correct if it occurred within one second of a significant deviation of a pointer; the subject was allowed to predict on the basis of prior observations, even though he was not observing that particular dial when he responded. Incentive payments were made on the basis of the proportion of correct responses during a session to the total number of significant deviations which occurred within that session.

The actual data of interest were movements and fixations of the subject's eyes during the trials. The eye positions were read from the film by two independent sets of readers in a manner analogous to the "Pilot Eye Movement Studies" (12).

3) Results : Table 1 shows the transition probabilities calculated on the basis of (11), and the observed transition probabilities. Fig. 1 shows the relation between signal bandwidth and obtained sampling frequencies.

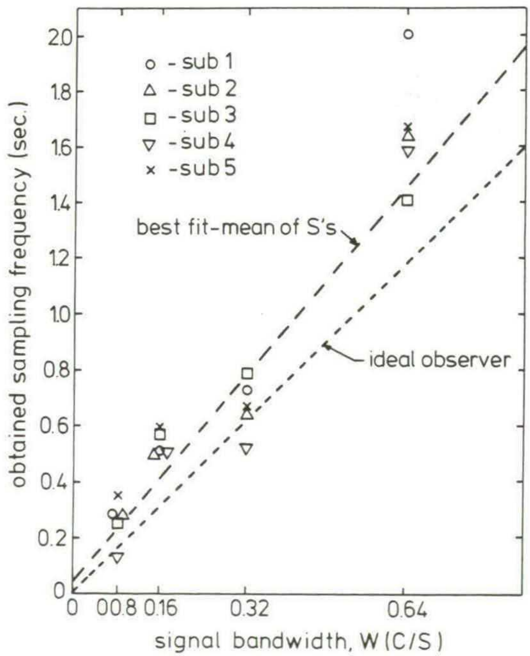


Fig. 1 Regression of obtained sampling frequencies on bandwidth.

TABLE 1

Obtained and Calculated Transition Probabilities.
The probability of transition between a and b is

$$P_{ab} = \frac{2P_a \; P_b}{\Sigma P^2} = 3.17 \; P_a P_b \tag{11}$$

Transition		Calculated Transition Probabilities	Obtained Transition Probabilities
P 4-2	(3.17) (0.537) (0.211)	0.361	0.324
P 4-1	(3.17) (0.537) (0.172)	0.293	0.297
P 4-1/2	(3.17) (0.537) (0.077)	0.134	0.133
P 2-1	(3.17) (0.214) (0.172)	0.117	0.112
P 2-1/2	(3.17) (0.214) (0.077)	0.052	0.051
P 1-1/2	(3.17) (0.172) (0.077)	0.012	0.040
		0.999	0.957

B. Discussion.

If the human observer behaved like an ideal observer, we should expect him to sample the displays at a rate of $2W$. Since we do not expect human observers to be either ideal or periodic, and since it is possible that periodic sampling cannot be achieved with this collection of four signal bandwidths, we expect and obtain a distribution of sampling rates. Since the observer is heavily loaded, and since the expected proportion of total time spent on an instrument should be proportional to the information flow rate from the instrument, we expect that the mean of the distribution of sampling rates will be some funtion of $2W$: $k \; 2W + c$ samples per unit time. *k* is the slope of the line relating predicted to obtained sampling rates; *c* is the *y* intercept or sampling rate for signals or instruments which do not in fact vary, or whose readings are in fact of no significance. Such fixations do occur and have been recorded in the "Pilot Eye Movement Studies". The data were originally recorded in the form of interobservation intervals, and have been converted into frequencies (Table 1) and plotted in Fig. 1. The mean best fitting line has a slope of 1.2 and an intercept of 0.04. However, if we had used in calculating our proportional frequencies or probabilities the observations of things other than the four instruments, we should almost certainty have found a greater intercept. Whether the same slope constant 1.2 will be found in general or whether it is specific to the task studied cannot be determined until further experiments are done with other task characteristics. In any event, the observed sampling frequencies are monotonically increasing with signal bandwidth. The transition probabilities predicted on the basis of (11) are remarkably close to those which were obtained experimentally. The correspondence suggests that the subjects were, on the average, distributing their samples in accord with the base probabilities that each instrument would be sampled. Of course this is tantamount to treating the man as a "Markov Processor". A later report will consifer the Markov model in detail.

IV. Conclusions

Thus, we can make a good prediction of link values on the assumption that an observer acts as a random sampling device constrained only by the base probabilities of observation of each of the things sampled. This does not mean that he is such a device, nor does it mean that the fine structure of his sampling possesses no discernible order .

Another set of data which allow a partial test of the adequacy of the random model for transitions was accumulated during the " Pilot Eye Movement Studies" (13). As indicated earlier, we have no record of the instrument readings during the flights. Therefore, we cannot test to see if pilots fixations are in accord with the model. We do have fixation frequencies, however, and by estimating the underlying relative frequencies of fixation, can calculate transition probabilities or link values on the basis of equation (11). This being done for some of the data of (13). The overall agreement is good with a correlation between predicted and obtained of greater than 0.9 (19)

There are at least three alternative behaviours which would produce the results we have obtained. It could be :

1) that human observers are in fact random sampling devices and nearly ideal observers; 2) that individual observers may have fixed patterns of scanning but these may differ among observers with the result that lumped statistics are in accord with the random model; 3) that individual observers may use fixed patterns of scanning for short periods and change the patterns from time to time. For the designer of instrument panels, the details of fine structure of behaviour are probably not too important unless there is reason to believe that relatively high work loadings may occur.

References

(1) H.M. James, N.B. Nichols, and R.S. Philips, "Theory of Servomechanisms," McGraw Hill book Co, Inc. , New York, N.Y. ; 1947.

(2) H.P. Birmingham and F.V. Taylor, "A Human Engineering Approach to the Design of Man-Operated Continuous Control Systems", Naval Research Lab., Washington, D.C. Rept. No. 4333; April, 1954.

(3) J.I. Elkind, "Characteristics of Simpel Manual Control Systems," M.I.T. Lincoln Lab., Lexington, Mass., Tech. Rept. No. 111; April 1965.

(4) --, private communication ; 1963

(5) R. Mayne, "Final Report : Human Dynamics Study," Goodyear Aircraft Corp., Akron, Ohio, Rept. No. GER-4750 ; April 1952.

(6) D.T. McRuer and E.S. Krendel, "Dynamic Response of Human Operators". Wright-Patterson Air Force Base, Ohio, Rept. No. WADC TR 56-524 ; October, 1957.

(7) J.D. North, "The Human Transfer Function in Servo Systems", in "Automatic and Manual Control", A. Tustin, Ed., Academic Press, Inc., New York, N.Y. ; 1952.

(8) G.A. Bekey, "The human operator as a sampled-data system", IRE TRANS. ON HUMAN FACTORS IN ELECTRONICS, vol. HFE-3, pp. 43-51; September, 1962.

(9) W.F. Battig, J.F. Voss, and W.J. Brogden, "Effect of frequency target intermittence upon tracking," *J. Exp. Psychol.*, vol. 49, pp. 49 ; April, 1955.

(10) C.A. Bennett, "Sampled-data tracking : sampling of the operator's output," *J. Exp. Psychol.*, pp. 51 ; 1956.

(11) J.W. Senders, " Tracking with Intermittent Stimuli (Forced Sampling), " Air Research and Dev. Command, Washington, D.C., Rept. No. ARDC TR-56-8 ; vol. 51 ; June, 1956.

(12) --, "Tracking with Intermittently Illuminated Displays," Wright-Patterson Air Force Base, Ohio, Rept. No. WADC TR-55-1978 (AD 80941) ; 1955.

(13) J.L. Milton, R.E. Jones, and P.M. Fitts, "Eye Fixations of Aircraft Pilots : II. Frequency, Duration and Sequence of Fixations When Flying the USAF Instrument Low Approach System (ILAS), "Wright-Patterson Air Force Base, Ohio Rept. No. USAF TR-5839 (ATI-65996) ; October, 1949.

(14) J.W. Senders, "Man's Capacity to Use Information from Complex Display," in "Information Theory in Psychology," H. Quastler, Ed., The Free Press, Glencoe, III ; 1955.

(15) C.E. Shannon and W. Weaver, "The Mathematical Theory of Communication," University of Illinois Press, Urbana, III,; 1949.

(16) W.E. Hick, "On the Rate of Gain of Information," Medical Research Council, Cambridge University, Cambridge, Engeland, A.P.U. 162/51; October, 1951.

(17) R. Hyman, "Stimulus information as a determinant of reaction time ," *J. Exp. Psychol.*, pp. 45 ; vol. 45 ; March , 1953.

(18) J.W. Senders, "Information input rates to human users : recent research results, " *WADC Symp. on AF Flight Instrumentation Program*, Wright-Patterson Air Force Base, Ohio ; 1958.

(19) J.W. Senders and K.A. Stevens, "Re-Analysis of the Pilot Eye-movement Data," Bolt Beranek and Newman, Inc., Rept. No. 1136, Cambridge, Mass. : May 1964.

3. *Experiments on Six-Instrument Scanning : General Features.*

These experiments were undertaken mostly as a replication, with a different number of dials, of the four instrument experiment. There was no intent to construct and test more elaborate models. The results of the earlier experiment had been so potentially useful that it was felt desirable to have a completely new replication with a different number of dials before putting the results to use. In consequence, no attempt was made to record the data which would have been useful in evaluating the Conditional Sampling Models. The cost of recording continuously the values presented on all the instruments and later correlating with the eye fixations to determine the value at the moment of fixation was deemed too great to be warranted given the unfinished state of the Conditional Models. However, it is possible to make some reasonable inferences about these latter models.

A number of experiments were done and some of these will be reported here. Some were failures. The Periodic Model assumes that the observer is engaged in sampling with the intent of reconstructing the signal. In most cases it is difficult if not impossible to have the subject do this. Further, it is assumed that the observer is reading the value of the signal presented on the instrument to some degree of accuracy which is or can be known. To have the subjects do this is perhaps impossible as well; The subjects do not read out the value as perceived nor would it be a reasonable task for them to undertake. The calling out of the values would certainly interfere with the monitoring behaviour itself and seriously alter it, if only because of the added workload associated with the verbalization. For these reasons, in particular, these experiments were done; they were an effort to work around these problems.

The purpose of one of the experiments was to explore the effects of correlations between the signals presented on different instruments to obtain a first estimate of the use that an observer could make of such correlations. The earlier four-instrument experiments had used completely uncorrelated signals, Gaussian and band-limited. These were generated by a random noise generator based on persumably random electron noise in a vacuum tube. Instruments in real machines present highly correlated signals in many, even most, cases. The physical nature of the machines makes this inevitable. It seemed reasonable to assume that if monitors of real machines could make use of correlations and coupling between signals, the workloads induced by monitoring would be less than those calculated on the basis of uncorrelated signals. As an example, it would be expected that a pilot would look at the altimeter if the rate of climb indicator showed a descent, rather than waiting until the altimeter came up in its normal turn.

Four experiments are reported here. These are as follows :

1. Remeasure the relation between signal bandwidth and frequency of fixation.

2. Measure the effect of a change in the composition of the signals driving the instruments from *pseudo-random* to *random*.

3. Measure the effect of intersignal correlation on fixational behaviour.

4. Measure the relation between signal bandwidth and frequency of fixation when the continious random signal has been converted into a discrete variable.

Method

All four experiments were conducted in identical environments. Each subject was seated on a laboratory chair of conventional height. A chin rest, adjustable in height, was provided. The combination of chair height and chin rest height was such as to place the eyes of the subject level with the center of the surface on which the meters were mounted. A camera was placed behind a hole at the center of the screen and aligned with the center of the subjects' eye-level. On the display surface were mounted micro-ammeters. Each meter scale covered the range from −50 through 0 to +50 micro-amperes. Three meters were above the camera and three below, aligned with the horizontal and spaced equidistant from one another. The instrument dials were 12 degrees apart at the viewing distance of 75 cm. Each meter was driven by a signal, as used by Elkind (6), composed of more than 40 sinusoids. The individual signals were pseudo-random, with a Gaussian amplitude density distribution. The signals were sharply band-limited by virtue of their construction, and were flat (had constant power) from some relatively low frequency to the cut-off frequencies listed.

Naturally there was not a continuous flat spectrum. Instead the components, closely spaced, were of the same power over the range used.

Except for experiment 3, the signal cut-off frequencies were .03, .05, .12, .20 , .32, and .48 hz. The sum of these bandwidths is 1.20 hz. and should represent a substantially full load for the monitor, on the assumption that at least two samples per cycle would have to be taken for effective monitoring, and that samples would last approximately .40 seconds. Experiment 2 used bandwidths of .02, .04, .08, .16, .32, and .64 hz. These sum to 1.26 hz. virtually the same.

One hour of recorded signals was available on magnetic tape. The complexity of the task was such that it was felt that no learning of the signals would occur in the course of the experiment. In general, each time the same signal was presented to the subject, the values seen would be different since the instants of sampling could hardly be the same as on any other presentation. The subjects were not aware that the signals were recorded and would be the same from trial to trial. In addition to these reasons for not expecting learning to the signals to occur, the precaution of using a different starting point each day was employed except where changes in level of performance were to be checked. Here the identical signal and starting point was used. All 5 subjects were run simultaneously although since only one camera was used, the moments of sampling were all different.

Subjects

The subjects were five high school students in the fourth year at Belmont High School. All were upper level students. Three were male and two female. All had adequate vision although it was not considered necessary that vision be adequate uncorrected, unless the use of eyeglasses interfered with the reading of the eye positions. The accompanying Figure 3 shows the general layout of the experiment. All the subjects were used in all the experiments.

Stimuli

The signals discussed above were fed to the micro-ammeters in series across the five subject positions. This ensured that each signal was presented to each subject with indentical current. For each subject position, the arrangement of the six bandwidths was different as to location on the panel, in a quasi-random manner.

Task

The task of the subjects was like that of the subjects in the four dial experiment. The subjects were asked to monitor the instrument and to report by pressing a handswitch whenever any of the pointers exceeded the 40 microampere level. The switches were silent so that the subjects would have no cue from one another as to the nature of the signals. Like the subjects in the earlier experiments, these subjects were told that they would receive a bonus based on how close their detection performance was to the actual number of times that the pointers exceeded $\pm$ 40 m.a. They were in fact rewarded in a random way. The size of the bonus, which varied randomly as well, was assigned randomly to the subjects and was not based on performance.* The subjects were given no further information about performance and no further instruction. They were seated in front of the panel and, at a signal, started to observe the instruments and to operate the handswitch provided. They performed their monitoring task for ten minutes and then received a rest of two minutes. This was repeated for one hour on each of ten days before data were actually taken. The choice of ten hours at one hour per day was made on the basis of the results of the four instrument experiment in which it was observed that performance was stabilized, as to frequency of fixation, after only 2 or 3 hours of exposure to the signals. Figure 3 shows schematically the experimental arrangement.

Recording

During the recording sessions the procedure was changed to allow time for the camera to be reloaded with film. Data were taken for II minutes and a 5 minute rest followed. A Bolex Reflex H-8 camera with an electric drive was used to record the eye positions. The camera operated at 12 frames per second, based on the internal governor of the camera. Calibration showed this to be accurate enough for the measurement task. One hundred foot reels of film were loaded into the camera and the motor was started when responses, and presumably monitoring, had begun. The reel lasted the II minutes which dictated the cycle time of the experiment. Each subject provided, therefore, about II minutes of data at the end of approximately 10 hours of monitoring behaviour. Since the subjects were photographed monitoring different portions of the recorded tapes, the eye movements reflect the responses to different pieces of the signals and may exhibit differences because of this fact.

The films were read on a Gerber Digital Data Reduction System, model GDDRS-3B in conjunction with a Gerber Scanner S-10-C and a Projector S-10-P. The system allowed direct conversion of distance to digital readout onto punched cards. The film was projected onto the surface of the analyzer screen and the hairline placed adjacent to the sprocket hole of the first frame in which the subject is looking at a particular instrument. The code for the direction of fixation was punched along with a numerical code for the frame number. Then the same sprocket for the last frame of that same fixation was similarly punched. Since the frame speed was fixed at 12 frames per second, the duration of the fixation, and the time of starting and stopping are known. The cards were then analyzed by conventional punched card readers and computational means.

* This decision was the result of the instrumental difficulties encountered and the belief that the performance would be unaffected by the policy.

Data Analysis

The objective of the experiments was to compare the predictions of the Simple Periodic Model with actual fixation behaviour. No more complicated analysis was contemplated. The work on the six dial experiments was done under contract with the N.A.S.A. and was effectively limited to the contractually stated goals. The means of frequency and duration of fixation were calculated from the film data and compared with the model predictions. At the time of the original work, no funds were available for extensive further analysis. Whether it would now be possible to obtain from the archives of N.A.S.A. the data tapes and whether it would be worth re-analysing the data, given that the definitive datum, the actual value of the signal at the moment of fixating, is missing from those tapes, is doubtful. The availability of improved eye position recording methods and micro-processors would make it more economical and useful to repeat the experiments.

Tables of data are presented with each of the four experiments. The data are plotted in Figures 4, 5 and 6, which appear at the end of Chapter III.

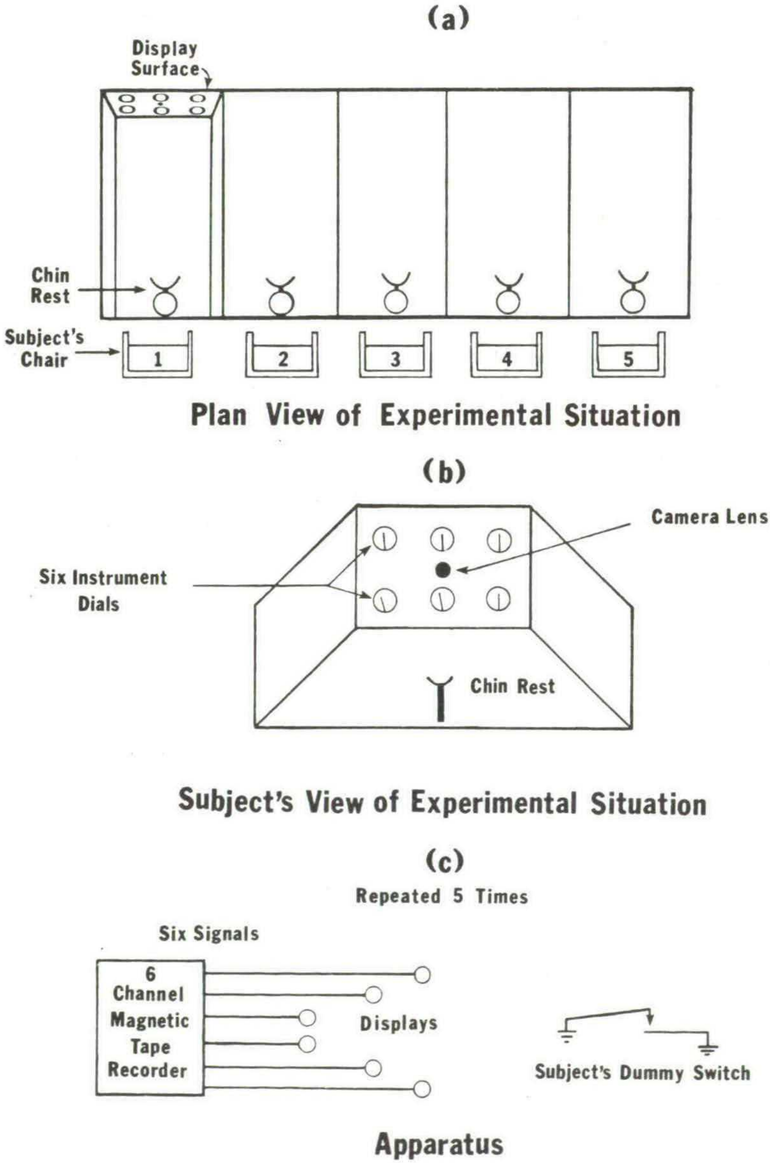

(a)
Display Surface
Chin Rest
Subject's Chair
1
2
3
4
5
Plan View of Experimental Situation
(b)
Camera Lens
Six Instrument Dials
Chin Rest
Subject's View of Experimental Situation
(c)
Repeated 5 Times
Six Signals
6 Channel Magnetic Tape Recorder
Displays
Subject's Dummy Switch
Apparatus

Figure 3

4. *Experiment 2 : The Relation between Signal Bandwidth and Frequency of Fixation.*

Method

Five subjects were run under the conditions described. Of these only three yielded useable data which could be subjected to analysis as planned. For one subject, one of the instruments jammed and failed to present the .48 hz. signal during the recording session. For another, the speed control of the recording camera failed and the films were taken at some higher but unknown rate. The data of the former failure provided an interesting and quite serendipitous experiment which will be discussed in another context. The data of the latter failure were discarded.

Approximately 33 minutes of behaviour of the three remaining subjects were studied. Three hundred feet of 8 millimeter film yielded approximately 24,000 frames of film. As the subjects made slightly more than 2 fixations per second, and there were 12 frames per second, about 4000 fixations were read and identified.

Results and Discussion

Frequency of Fixation

The frequency of fixation on each of the six signals for each of the three subjects is given in Table 1a and plotted as a function of signal bandwidth on Figure 4a. The same data modified in accord with the RCM are given in Table 1b. Although the linear correlation is high, .989, the data are curvilinear in the form predicted by the RSM and by none of the others.

The models elaborated in Chapter II fall into two classes : that which admits forgetting and those which do not. Those which do not, predict that the relationship between sampling frequency and bandwidth should be linear and should pass through the origin. It is clear that the experimental data intersect the ordinate at about .2 per second. Even for the lowest bandwidth a sample is taken approximately every 5 seconds. Such a result is compatible only with a model which attributes to the observer some forgetting between fixations.

TABLE 1a

Frequency of fixation (per second) as a function of signal bandwidth (Raw Data)

Best Fitting Line to the means : $\overline{F} = .21 + .63w$; $r = .989$

Bandwidth hz.	.48	.32	.20	.12	.05	.03	sum F/S
Subject no :							
2	.495	.448	.377	.298	.252	.172	2.04
3	.505	.393	.302	.320	.239	.278	2.04
5	.483	.463	.379	.235	.218	.196	1.97
Mean	.494	.435	.353	.284	.236	.215	2.02

TABLE 1b

Frequency of fixation (per second) as a function of signal bandwidth (in accord with RCM)

Best Fitting Line to the means : $\overline{F} = .17 + 1.34\,w$; $r = .999$

Bandwidth hz.	.48	.32	.20	.12	.05	.03	sum F/S
Subject no :							
2	.825	.611	.453	.331	.262	.176	2.58
3	.842	.536	.363	.356	.249	.285	2.67
5	.805	.632	.455	.261	.227	.201	2.56
Mean	.824	.593	.424	.316	.246	.221	2.60

Thus it is clear that the PSM fails for these data. The data as recalculated in accord with the RCM are closer to the theoretical line but only at the bandwidths at and above .2 hz. bandwidth. Below the .2 bandwidth, the data lie above the model prediction and increasingly so with decreasing bandwidth. Because of the flattening of the data for the lower bandwidths the slope of the line of best fit by least squares is very low, .63.

The slope of the line of best fit to the RCM recalculated data is higher at 1.34 but this is still very much below the expected slope which is nearly equal to 2.0.

It is evident that the instrument monitors do not behave like signal reconstructors or random choosers from an urn and with perfect memory. We should not be surprised at this result were it not for fact that the results of the Four Instrument Scanning Experiment so strongly supported the PSM. There are many possible sources of difference. Among these is the make-up of the signals used to drive the instruments. The next experiment attacks this question. A comprehensive discussion of the results will be presented after each of the experiments has been presented and briefly discussed.

Duration of Fixation.

Although the principal purpose of the experiment was to relate frequency of fixation to signal bandwidth, duration data were also gathered.

The duration of fixation on each of the six signals for each of three subjects is given in Table 2a and plotted as a function of signal bandwidth on Figure 5a. The same data modified in accord with the RCM are given in Table 2b. The results for the raw data are in general accord with expectation, the slope being .08. The data as recalculated in accord with the RCM have a negative slope. The meters were all identical with respect to required precision of reading and the durations of observation should be the same for all. In both cases the duration at intercept is .45 seconds. This value is close to the mean of data gathered by Fitts and his colleagues in actual flight. (1)

TABLE 2a

Duration of fixation (seconds) as a function of signal bandwidth (Raw Data)

Weighted $\overline{D}$ = .47 sec.

Bandwidth hz.	.48	.32	.20	.12	.05	.03
Subject no :						
2	.58	.43	.44	.49	.49	.43
3	.48	.44	.53	.45	.47	.38
5	.42	.50	.49	.45	.48	.44
Mean	.49	.46	.49	.46	.48	.42

TABLE 2b

Duration of fixation (seconds as a function of signal bandwidth (in accord with RCM)

Bandwidth hz.	.48	.32	.20	.12	.05	.03
Subject no :						
2	.35	.32	.37	.44	.47	.42
3	.29	.32	.44	.41	.45	.37
5	.25	.37	.41	.41	.46	.43
Mean	.29	.34	.41	.41	.46	.41

Allocation of Time

The percent of total time spent on each signal is linear with bandwidth. Table 3 and Figure 6a present the data. It is evident that the information generation rate of the presented signal is a powerful determinant of the allocation of visual attention. The intercept of .09, which is in a sense lost time, may represent the time taken by eye movements themselves.

In addition to the time lost in eye movements, there are about 5 % of the total number of frames which cannot be assigned to any of the instruments. For this reason the sum of the times allocated to the various signals add up, generally, to less than 1.0.

TABLE 3

Proportion of total time allocated to signals as a function of bandwidth

$$\% T = .093 + .326 W \; ; r = .986$$

Bandwidth hz.	.48	.32	.20	.12	.05	.03	sum
Subject no :							
2	.287	.193	.166	.146	.123	.074	.989
3	.242	.173	.160	.144	.112	.106	.937
5	.203	.232	.186	.106	.105	.086	.918
Mean	.242	.200	.173	.131	.113	.090	.949

5. Experiment 3 : The Effect of Signal Composition on Fixational Behaviour.

Rationale

The signals used in experiments 2, 4 and 5 consisted of pseudo-random time functions composed of the sum of a very large number of sine waves recorded onto the same tape track. The frequencies and phases of the components were chosen so as to yield a very long time between repetitions of the exact sequence of events. Such signals have an amplitude density distribution that is very nearly Gaussian. They have been used with apparent great success in the investigation of tracking behaviour . (6)

The subjects did not consider these signals to have the same quality of randomness which other signals characterized as "random" have. In general, there was an apparent pendulosity which made the signals relatively predictable. The lack of exceedingly low frequencies was the cause of this appearance. If the signals did have serial redundancy, the rate of increase of uncertainty would be much smaller than in the case of truly random time functions, and the permissible interval between samples would be longer by the same argument. The subjects would, in a sense, be underworked, rather than working very nearly at their limit. The data from Experiment 2 showed that the subjects characteristically spent some (from 2 % to 8 %) of their time looking at places other than the instruments, depending on the observer and the experiment. Since the calculated workload was very close to 100 %, there was a reasonable suspicion that the signal characteristics influenced the subjects in an unexpected way.

Method

Experiment 2 was repeated using signals generated by a "random noise generator". Zener diode noise was produced and filtered by a three section filter to give a cutoff of 18 db. per octave and then recorded on a six-channel Mnemotron tape recorder. The various frequencies were produced by changing the speed of the tape recorder.

Data were gathered on two subjects from the group used in Experiment 2. It is of interest to note that although the calculated sum of the bandwidths was only .06 hz. greater than in Experiment 2, the subjects reported that the signals were much 'harder" to monitor, and expressed themselves as being more fatigued after one hour of these new signals than they had been with the old. The films yielded 15,832 frames of data. These represented 2,800 fixations. The data were gathered after 5 hours of monitoring of the new signals to allow learning to occur and performance to stabilize.

Results and Discussion

Frequency of Fixation

The frequency of fixation on each of the six instruments is shown in Table 4a and 4b and in Figure 4b.

The slope constants for both the raw data and the data recalculated in accord with the RCM are not significantly different from those found in Experiment 2. The data exhibit higher than predicted sampling rates for the low frequencies and lower than predicted sampling rates for the high frequencies. These results are similar to those found in Experiment 2. The curvilinear relationship between fixation frequency and signal bandwidth is more pronounced in this case presumably due to the fact that the bandwidths extend to .64 hz. instead of .48 hz. as in the former experiment.

The difference in signal composition made no appreciable difference in the experimental results. The comparison of the two fixation frequencies on the .32 hz. signal common to both sets shows that the difference is not significant (t is less than 1.0).

Again those models which predict a linear relationship passing through the origin fail. The data are distincly non-linear. It is conceivable that with some sufficiently low signal bandwidth the fixation frequency would approach zero but these data give no indication of when that might occur.

TABLE 4a

Frequency of Fixation (per second) as a Function of Signal Bandwidth (Raw Data)

Best Fitting Line for Means : $\overline{F} = .221 + .605w$; $r = .959$

Bandwidth hz.	.64	.32	.16	.08	.04	.02	sum F/S
Subject no :							
1	.504	.502	.265	.303	.215	.233	2.02
2	.639	.445	.423	.253	.263	.127	2.15
Mean	.572	.478	.344	.278	.239	.180	2.09

TABLE 4b

Frequency of Fixation (per second) as a Function of Signal Bandwidth (in accord with RCM)

Best Fitting Line for Means : $\overline{F} = .162 + 1.53w$; $r = .999$

Bandwidth hz.	.64	.32	.16	.08	.04	.02	sum F/S
Subject no :							
1	1.024	.673	.297	.323	.222	.237	2.78
2	1.298	.587	.485	.270	.272	.129	3.04
Mean	1.161	.635	.391	.297	.247	.183	2.91

Duration of Fixation

The duration of fixation on each of the six instruments is shown in Table 5a and 5b. The duration of fixation on the .64 hz. instrument is more than nine standard deviations away from the mean of the other five durations. When the data are recalculated in accord with the RCM the mean durations are homogeneous and none of them is significantly distant from the others. The result is what one would expect for instruments which have identical requirements for precision of reading.

Again the data demonstrate an increase of time spent on an instrument with an increase of the bandwidth of the signal driving that instrument. The slope is larger but not significantly so (t is less than 1.0). The data are very nearly linear with bandwidth. Once again we find that the time spent is a linear function of signal bandwidth despite the non-linearity of the fixation frequency relationship. This is accomplished by a reciprocal non-linearity of the relation between duration and bandwidth. The intercept of .07 is not significantly different from the value of .09 found in Experiment 2.

TABLE 5a

Duration of Fixation (seconds) as a function of signal bandwidth (Raw Data)

Weighted $\overline{D}$ = .48 sec.

Bandwidth hz.	.64	.32	.16	.08	.04	.02
Subject no :						
1	.73	.39	.43	.39	.41	.47
2	.63	.45	.39	.37	.33	.41
Mean	.68	.42	.41	.38	.37	.44

TABLE 5b

Duration of Fixation (seconds as a function of signal bandwidth (in accord with RCM)

Bandwidth hz.	.64	.32	.16	.08	.04	.02
Subject no :						
1	.36	.29	.38	.37	.40	.46
2	.31	.33	.34	.34	.32	.40
Mean	.33	.31	.36	.35	.36	.43

TABLE 6

Proportion of total time allocated to signals as a function of bandwidth

$$\% T = .07 + .49\,w \; ; r = .995$$

Bandwidth hz.	.64	.32	.16	.08	.04	.02
Subject no :						
1	.37	.20	.11	.12	.09	.11
2	.40	.20	.16	.09	.09	.05
Mean	.39	.20	.14	.11	.09	.08

6. *Experiment 4 : The Effect of Intersignal Correlations on Fixational Behaviour.*

Rationale.

In virtually any operational system, the various signals presented to the human observer are not statistically independent. This is the result of the inherent physical coherence of such systems. In an aircraft, for example, many of the instruments present data which are the derivative or the integral of the data presented on some other instrument. Rate of climb is the derivative of altitude; rate of turn is the derivative of heading. In other cases, the same data are presented in different ways. The physical nature of machines demands that the various indices of its performance be related since the underlying processes of the machine are related as well. Engine thrust, rate of climb, acceleration, altitude, and speed, and angle of attack all interact and their representations on instruments combine to present a picture of the machine's state to the pilot.

In the laboratory, on the other hand, there is usually a deliberate effort to present signals which are completely unrelated. The experiments just reported are quite typical in this regard.

Do human operators take advantage of the inter-signal correlations in order to reduce their information processing load? If we are to be able to apply any of the scanning models it would be helpful to know the answer to this question. The experiment here reported was an attempt to investigate this aspect of visual sampling.

The realistic way to approach the problem would be to construct a system which had physical coherence. A simulated aircraft would serve. Within the context of a set of laboratory experiments, however, it is preferable to maintain the non-coherence and independence of displayed signals, one from the other, and use purely statistical correlation. The further advantage of this approach is that the degree of relationship between signals can be easily manipulated by the experimenter. The experiment altered the relationship of one pair of signals. To have presented more than one correlated pair would have made it difficult to sort out the effects, if any, since the observer has only so much time and is therefore unable to change behaviour toward one instrument without altering it toward all the others to some extent. The method chosen was, therefore, to mix signals driving two of the instruments. The nominal .12 hz. display was used because the sampling frequency toward this signal by the subjects in Experiment 2 was very nearly on the function predicted by the model. We might expect that the behaviour toward this signal would be more sensitive than a lower bandwidth, and less variable than that toward a higher one. The signalș from one source was mixed with that from another so that a correlation from -1.0 to +1.0 could be had at will. For example if .707 of signal *a* is added to .707 of signal *b* , which is equal in power to signal *a*, the composite signal has a power equal in power to *a* or *b* and has a correlation of .50 with either of them.

Method

The same instrument display was used and a pilot study run with the values of .707 as above. The particular composite was a combination of .12 hz. and .20 hz. The .20 hz. signal was also presented. The composite was correlated .5 with the .20 hz. The behaviour toward this composite was not distinguishable from that toward the .12 hz. signal in Experiment 2. In the experiment itself a higher value of correlation, .81 was chosen. The composite signal was a mixture of .90 of the .20 hz. signal and .44 of the .12 hz. signal.

The same five subjects used in experiment 2 were used in this experiment and were trained on the array of signals including the composite. They were not informed of the existence of the correlation between two of the signals.

Approximately 42,000 frames of film were read. These yielded data from approximately 3,500 seconds of behaviour.

Results and Discussion

Tables 7a and b, 8a and 8b, and 9 and Figures 4c, 5c and 6c present the data for frequency of fixation, durations of fixation, and percent of total time spent, all as functions of signal bandwidth.

Only one subject commented on the existence of the correlation between the signals presented on two of the instruments.

Frequency of Fixation.

The effect of the experimental manipulations was small if indeed there was any. The frequency of fixation on the instrument driven by the mixed signal was essentially the same as if the indicator had been driven by a .12 hz. signal. The duration of observation was, however, very much lower than the weighted mean duration and this is reflected in a low percentage of time spent on that indicator.

TABLE 7a

Frequency of fixation (per second) as a function of signal bandwidth (Raw Data)

Best Fitting Line : $\overline{F} = .15 + .56\,w \;;\; r \doteq .973$

Bandwidth hz.	.48	.32	.20	$\frac{.12}{.20}$	.05	.03	sum F/S
Subject no :							
1	.363	.237	.254	.160	.175	.206	1.395
2	.337	.340	.334	.175	.202	.166	1.554
3	.372	.326	.202	.220	.083	.085	1.288
4	.553	.451	.462	.363	.172	.095	2.096
5	.360	.385	.252	.187	.176	.249	1.609
Mean	.397	.348	.301	.221	.162	.160	1.589

Again the fixation frequencies are lower than those predicted by the PSM. The slope of the function for the raw data is .61, essentially the same as for the two earlier experiments. The correlation between frequency of fixation and bandwidth,.973 , is very much less for the raw data than for the RCM, which has a correlation of .999.

TABLE 7b

Frequency of fixation (per second) as a function of signal bandwidth (in accord with RCM)

Best Fitting Line : $\overline{F} = .12 + 1.13w$; r = .999

Bandwidth hz.	.48	.32	.20	$\frac{.12}{.20}$	.05	.03
Subject no :						
1	.61	.32	.30	.18	.18	.21
2	.56	.46	.40	.19	.21	.17
3	.62	.44	.24	.24	.09	.09
4	.92	.62	.55	.40	.18	.10
5	.60	.53	.30	.21	.18	.26
Mean	.66	.47	.35	.24	.17	.17

Duration of Fixation

The duration of fixation on the composite signal was significantly less than the mean of the durations on the other four signals ($t = 3.6$; 2df.). The subjects were using shorter fixations, and the experimental hypothesis was confirmed. The correlation between the two related signals was large. Whether, with sufficient training smaller values of correlation would have an effect is a question to be answered by later experimentation. There is no theoretical basis for calculating a threshold correlation above which an effect would be elicited.

The duration of fixation on the .20 hz. signal was significantly greater than the mean of the durations on the other four signals ($t = 3.9$; 2 df.) This result is not explicable by the models. There is no compensation in the form of a reduction of frequency of fixation on the .20 hz. signal. As a result there is an increase in the amount of time allocated to that signal.

Allocation of Time.

As indicated above, the time spent on the .20 hz. signal is larger than in any of the other experiments. This is unexpected. We also find the expected reduction in time spent on the composite signal. The slope of .467 is between the slopes found in the two earlier experiments.

TABLE 8a

Duration of fixation (seconds) as a function of signals bandwidth (Raw Data)

Weighted $\overline{D}$ = .62 sec.

Bandwidth hz.	.48	.32	.20	$\frac{.12}{.20}$	.05	.03
Subject no :						
1	.71	.86	.77	.49	.58	.49
2	.62	.72	.73	.44	.50	.48
3	.77	.74	.99	.50	.70	.54
4	.47	.47	.59	.32	.41	.47
5	.70	.63	.70	.43	.51	.52
Mean	.65	.68	.76	.44	.54	.50

TABLE 8b

Duration of fixation (seconds) as a function of signal bandwidth (in accord with RCM)

Bandwidth hz.	.48	.32	.20	$\frac{.12}{.20}$	.05	.03
Subject no :						
1	.43	.63	.64	.44	.56	.48
2	.37	53	.61	.40	.48	.47
3	.46	.54	.83	.45	.67	.53
4	.28	.34	.49	.29	.39	.46
5	.42	.46	.98	.39	.52	.51
Mean	.39	.50	.63	.39	.52	.49

TABLE 9

Proportion of total time allocated to signals as a function of bandwidth

$T = .07 + .43\,w \;;\; r = .90$

Bandwidth hz.	.48	.32	.20	$\frac{.12}{.20}$	.05	.03
Subject no :						
1	.26	.20	.20	.08	.10	.10
2	.21	.24	.24	.08	.10	.08
3	.29	.24	.20	.11	.06	.05
4	.26	.21	.27	.12	.07	.04
5	.25	.24	.18	.08	.09	.13
Mean	.25	.23	.22	.09	.08	.08

7. *Experiment 5 : The Effects of Dichotomizing the Display of a Continuous Signal on Fixational Behaviour*

The theory and the models which derive from it are based on the idea that it is the nature of the signal driving the display that determines the fixational behaviour. There has been no attempt to incorporate any function of the design of the display instrument into the predictions of frequency and duration of fixation. It is clear, however, that some extreme variations from good design practice must make a difference in the behaviour of the obderver. Thus if the display is badly illuminated, very much too small, nonlinear or has any of a very large number of possible distortions of its readings, the time taken by an observer to extract information from it will be increased and the frequency of reading errors will also be increased. The result of this latter will almost surely be an increase in the frequency of fixation. Thus, we would expect that for a badly designed display there will be more time in general spent in fixating it, and that this will be the result of both longer fixations and more of them.

Many information displays in modern systems are in the form of warnings. The signal does not appear until some set-point or limit has been reached. Since the models discussed here are all concerned with internally generated attentional processes, there would be no reason to expect that any of them would be of value for the predicting of the fixational behaviour, if indeed there were any, toward a warning light. As an alternative, it was decided to use a "flag" which could not be seen unless it were fixated.

Hypotheses

There are two quite opposite hypothesis. One is that if the observers, when the entire signal is presented to them, make use of the derivative information in estimating the probability that the signal will exceed the critical level, then the change of display form would eliminate this information and make it necessary for them to sample the display more often than would otherwise be the case. A second hypothesis, and one that is equally tenable, is that if an observer normally takes account only of the position of the indicator and notes that it is above or below the critical level without regard for its velocity, then the transformation of the signal into 0's and 1's will facilitate his observational process and require of him fewer observations than would otherwise be the case. One would also expect a reduction in the duration of fixation in both cases as compared with the continuous display of the signal since all decisions are of one bit only. It is possible, of course, that some observers deal with such signals in one way, and others in the other way. Further, if the mean durations of the large deviation signals were short, then there would be a need for the monitor to look more frequently than if they were long. This is, of course, merely another way of saying that high bandwidth signals must be looked at more often than low bandwidth signals.

Procedure and Subjects

The experimental procedure was the same as earlier described. The subjects were those used in the earlier experiments and had become well practiced in the monitoring of the six dial array. The signal bandwidths were the same as those used before : .03, .05, .12, .20, .32, .48 hz. in bandwidth. The .12 hz. signal was chosen for conversion for the reasons given earlier.

The signals with larger bandwidths were already being fixated with very high frequencies and there might have been little room for important change except at the expense of the other signals. The lesser bandwidths were already being sampled at rates which seemed much in excess of what should be the case on any of the models. Previous fixational behaviour toward the .12 hz. bandwidth signal had been consistent, lying on or close to the theoretical values, and it might be expected that departures from the expected values would be easier to detect since there would be more scope for them to occur.

The subjects were presented with the new display and the operation was explained to them. It was emphasized that the new signal display required the same detection task as had previously been the case. After five days of practice, one hour per day, data were taken for ten minutes on each of the subjects. Data were read from 40,204 frames of film which resulted in a total of 6,100 fixations.

Although there would undoubtedly exist brief carry-over from the previous experiments which these same subjects had been engaged in, the large training time on the simple variation was deemed sufficient to establish new and appropriate behaviour toward the one signal which had been altered. When we take subjects into the laboratory to study their eye-movements toward dynamic stimuli, we are using subjects who have already lived some twenty years or so. If they were awake two-thirds of the time, and if they made two eye movements per second they would, by the time of the experiment have made some 2 x 109 eye-movements to a wide variety of stimuli. The brief experience in the laboratory would merely be one more occasion on which some new and appropriate adaptation must be made. Relatively rapid shifting of the distribution of fixations and of the statistics of fixations in accord with circumstances must have become part of the heritage of human beings. We should not expect, therefore, to have important effects from the earlier experiments appearing here.

Results and Discussion

Tables 10, 11 and 12 present the frequency of fixation, the duration of fixation and the allocation of time, all as functions of the bandwidth of the signal driving the instruments Figures 4d, 5d and 6d show the means for all three variables as functions of bandwidth.

Frequency of Fixation

The mean frequency of fixation on the flag signal is less than on the signals with bandwidths immediately higher and lower. It is markedly different from the behaviour toward the .12 hz. signal in the earlier experiments. Otherwise the data are, overall, very similar to those of the earlier experiments. The low bandwidths are sampled more often than any of the linear models predicts and the high bandwidths are sampled less often. The slope of the best fitting line is .63. This value is essentially the same as those of the earlier experiments.

Duration of Fixation

The durations of fixation are approximately like those found in the earlier experiments The fixation duration on the dichotomized signal is not significantly different from the durations of fixation on the other signals. If the display method had any effect, it was on the frequency of fixation.

TABLE 10a

Frequency of fixation (per second) as a function of signal bandwidth (Raw Data)

Best Fitting Line : $\overline{F} = .18 + .63\,W\ ;\ r = .97$

Bandwidth hz.	.48	.32	.20	.12	.05	.03	sum F/S
Subject no :							
1	.412	.350	.343	.128	.262	.245	1.74
2	.579	.469	.279	.116	.208	.191	1.842
3	.532	.339	.334	.411	.273	.263	2.152
4	.438	.488	.156	.175	.204	.063	1.524
5	.413	.410	.269	.285	.242	.233	1.852
Mean	.475	.411	.278	.223	.238	.199	1.824

TABLE 10b

Frequency of fixation (per second) as a function of signal bandwidth (in accord with RCM)

Best Fitting Line : $\overline{F} = .14 + 1.31w$; $r = .98$

Bandwidth hz.	.48	.32	.20	.12	.05	.03
Subject no :						
1	.69	.48	.41	.14	.27	.25
2	.97	.64	.33	.13	.22	.20
3	.89	.46	.40	.46	.28	.27
4	.73	.67	.19	.19	.21	.06
5	.69	.56	.32	.32	.25	.24
Mean	.79	.56	.33	.25	.25	.20

TABLE 11a

Duration of fixation (seconds) as a function of signal bandwidth (Raw Data)

Weighted $\overline{D}$ = .54 sec.

Bandwidth hz.	.48	.32	.20	.12	.05	.03
Subject no :						
1	.72	.59	.56	.44	.54	.40
2	.68	.48	.35	.55	.40	.41
3	.42	.54	.54	.34	.42	.51
4	.81	.65	.44	.45	.55	.42
5	.55	.48	.64	.55	.49	.42
Mean	.64	.55	.51	.46	.48	.44

TABLE 11b

Duration of fixation (seconds) as a function of signal bandwidth (in accord with RCM)

Bandwidth hz.	.48	.32	.20	.12	.05	.03
Subject no :						
1	.43	.43	.47	.40	.52	.39
2	.41	.35	.29	.50	.38	.40
3	.25	.40	.45	.31	.40	.50
4	.49	.48	.37	.41	.53	.41
5	.33	.35	.53	.50	.47	.41
Mean	.38	.40	.42	.42	.46	.42

Allocation of Time

Again the data are similar to those of the other experiments. The best-fitting line has a slope of .48 and an intercept of .07. The correlation between bandwidth and percent time allocated is .98.

TABLE 12

Proportion of total time allocated to signals as a function of bandwidth

$$\% T = .07 + .48W \; ; \; r = .98$$

Bandwidth hz.	.48	.32	.20	.12	.05	.03
Subject no. :						
1	.30	.21	.10	.06	.14	.10
2	.39	.23	.1	.06	.11	.1
3	.22	.18	.19	.14	.11	.13
4	.35	.31	.06	.08	.11	.03
5	.23	.20	.17	.16	.12	.10
Mean	.30	.23	.14	.10	.11	.09

8. *General Discussion of the Six-Instrument Experiments*

The four experiments yield similar and consistent results. In each case the slope of the best-fitting line is about .65 for the raw data and about 1.3 for the data recalculated in accord with the assumptions of the RCM. In no case does the slope approach the values found for the subjects in the four-dial experiments. There are a number of differences between the experiments. The most evident is, of course, the number of dials to be monitored by the subjects. The subjects in the four-dial experiments were undergraduates in a college; the subjects in the six-dial experiments were high school students. The signals in the four-dial experiments were random with continuous spectra; those in the six-dial experiments were (with the exception of the second study reported) composed of many sine functions added together. The "bonuses" in the four-dial experiments were based on actual counts of successful detection of the critical signals; those in the six-dial experiments were based on a contrived random reward program unrelated to actual performance. The frequencies used in the four-dial experiment were .08, .16, .32 and .64 hz. Those used in the six-dial experiment were .03, .05, .12, .20, .32 and .48 (except for the second experiment reported which used .02, .04, .08, .16, .32, and .64.) In all cases the sum of the bandwidths, which should be an index of the total demand of the multi-dial task, was very nearly the same : 1.2, 1.2, and 1,26 hz.

In none of the six-dial experiments did any of the subjects scan in accord with the Periodic Model. In the four-dial experiments, the subjects generally yielded data which lay above the line of the Periodic Model. Which of the many differences was responsible, or what combination was responsible cannot be answered by the data of these experiments. The four-and six-dial studies were done many years apart and it had not been considered necessary to examine the earlier results by exact replication.

Since the four-dial results lay above the Periodic Model prediction and the six-dial results below, the differences in the two experimental series were potent. One thing is certain: the Conditional Sampling Models make predictions that are very far away from the performance of the subjects in the six-dial experiments. How can such data be appropriate for a task like that being performed by the subjects in these later experiments ? One way, of course, is for the velocities of the pointers on the dials to exceed the threshold for velocity perception so that the number of samples to be taken for signal specification is reduced. If that were the case one would expect to see a longer fixation time and this does not appear. In both sets of experiments the duration of fixation was about .45 seconds.

Among the differences between the four-dial and the six-dial experiments was the difference in reward basis. In one the reward was truly based on performance; in the other it was simulated. We have no way of knowing, within the confines of the experiments and the data, whether this difference was responsible for the difference in behaviour.

The age differences and the differences in amount of schooling should not have been responsible for the differences in the results. The high school students in the six-dial experiments were in the college stream and were presumably of the same intellectual sort as those who had already entered college. The composition of the signals to be monitored was tested as a source and found wanting. The subjects' reports were that there were differences in the signals but whatever the differences may have been there was no discernable effect on the performances.

The increase in number from four to six is great only on a percentage basis, yet it is conceivable that this change was at the root of the change in experimental outcome. The introduction of the two very low bandwidth signals could have created an overload situation (because of short term memory loss for the low bandwidth signals) which resulted in a reduction of time to be allocated to the higher bandwidth signals.

Overall there is great consistency in the data. The frequencies of fixation for each of the bandwidths are similar; the best-fitting lines have similar intercepts and slopes. Whatever the subjects are doing, they are doing it consistently. Where there were deviations in the frequency of fixation, there was, for the most part, compensation in the durations, so that the time allocated was nearly linear with bandwidth.

The information processing model of the human operator is again sustained; it is the information processing rate of the signal which appears to be the principal determiner of the fixation time allocated to a signal to be monitored.

All the models predict a monotone function of frequency of fixation and allocation of time as a function of bandwidth. Only the model which incorporates both the uncertainty (due to short term memory loss at low bandwidths) and the derivative sampling (at high bandwidths) predicts the actual form of the data which was found in all the experiments.

The aggregate data from experiments 1, 3 and 4 are presented in Tables 13a and b, 14a and b, and 15. The corresponding functions are shown in Figures 4e, 5e and 6e. In Figure 6e the data from experiment 2 are indicated by the *x*'s. It can be seen that the data from the two different kinds of signal, and two different sets of bandwidths are remarkably close to a common line function.

TABLE 13a

Frequency of fixation (per second) as a function of signal bandwidth (Raw Data)

Best Fitting Line : $\overline{F} = .180 + .61\,w \; ; \; r = .990$

Bandwidth hz.	.03	.05	.12	.20	.32	.48
Exp. 1	.215	.236	.284	.353	.435	.494
Exp. 3	.160	.162	.221	.301	.348	.397
Exp. 4	.199	.238	.223	.278	.411	.479
Mean	.191	.212	.243	.311	.398	.455

TABLE 13b

Frequency of fixation (per second) as a function of signal bandwidth (in accord with RCM)

Best Fitting Line . $\overline{F} = .14 + 1.26\,w \; ; \; r = .997$

Bandwidth hz.	.03	.05	.12	.20	.32	.48
Exp. 1	.221	.246	.316	.424	.593	.824
Exp. 3	.164	.169	.246	.363	.477	.662
Exp. 4	.204	.249	.248	.335	.563	.792
Mean	.196	.221	.270	.374	.544	.759

TABLE 14a

Duration of fixation (seconds) as a function of signal bandwidth (Raw Data)

Weighted $\overline{D} = .54$

Bandwidth hz.	.03	.05	.12	.20	.32	.48
Exp. 1	.42	.48	.46	.49	.46	.49
Exp. 3	.50	.54	.44	.76	.68	.65
Exp. 4	.44	.48	.47	.51	.57	.64
Mean	.45	.50	.46	.59	.57	.59

TABLE 14b

Duration of fixation (seconds) as a function of signal bandwidth (in accord with RCM)

Bandwidth hz.	.03	.05	.12	.20	.32	.48
Exp. 1	.42	.47	.44	.37	.32	.35
Exp. 2	.49	.52	.37	.63	.50	.39
Exp. 3	.43	.46	.42	.42	.42	.38
Mean	.45	.48	.41	.47	.41	.37

TABLE 15

Proportion of total time allocated to signals as a function of bandwidth

$T = .08 + .41\,W \; ; \; r = .982$

Bandwith hz.	.03	.05	.12	.20	.32	.48
Exp. 1	.09	.11	.13	.17	.20	.48
Exp. 3	.08	.09	.10	.23	.24	.26
Exp. 4	.09	.12	.10	.14	.23	.30
Mean	.09	.11	.11	.18	.22	.27

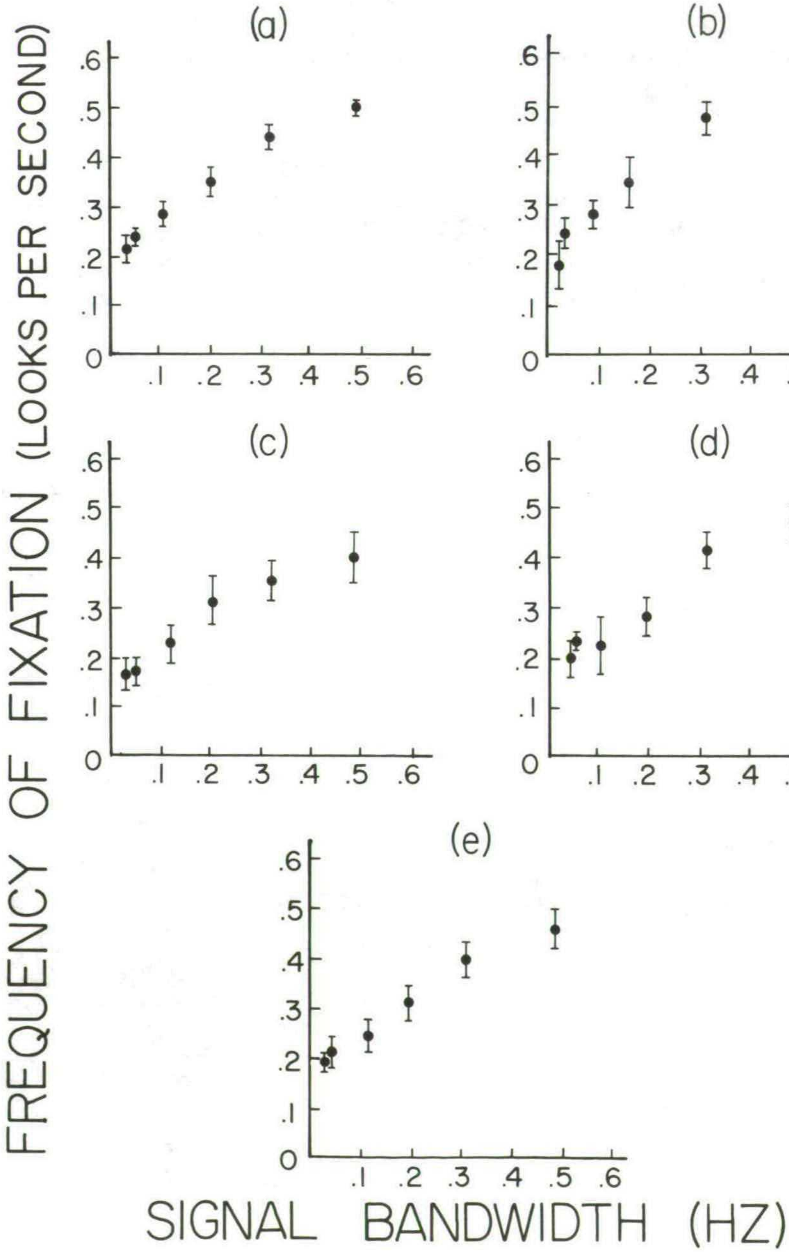

FREQUENCY OF FIXATION (LOOKS PER SECOND)
SIGNAL BANDWIDTH (HZ)
(a)
(b)
(c)
(d)
(e)

Figure 4.

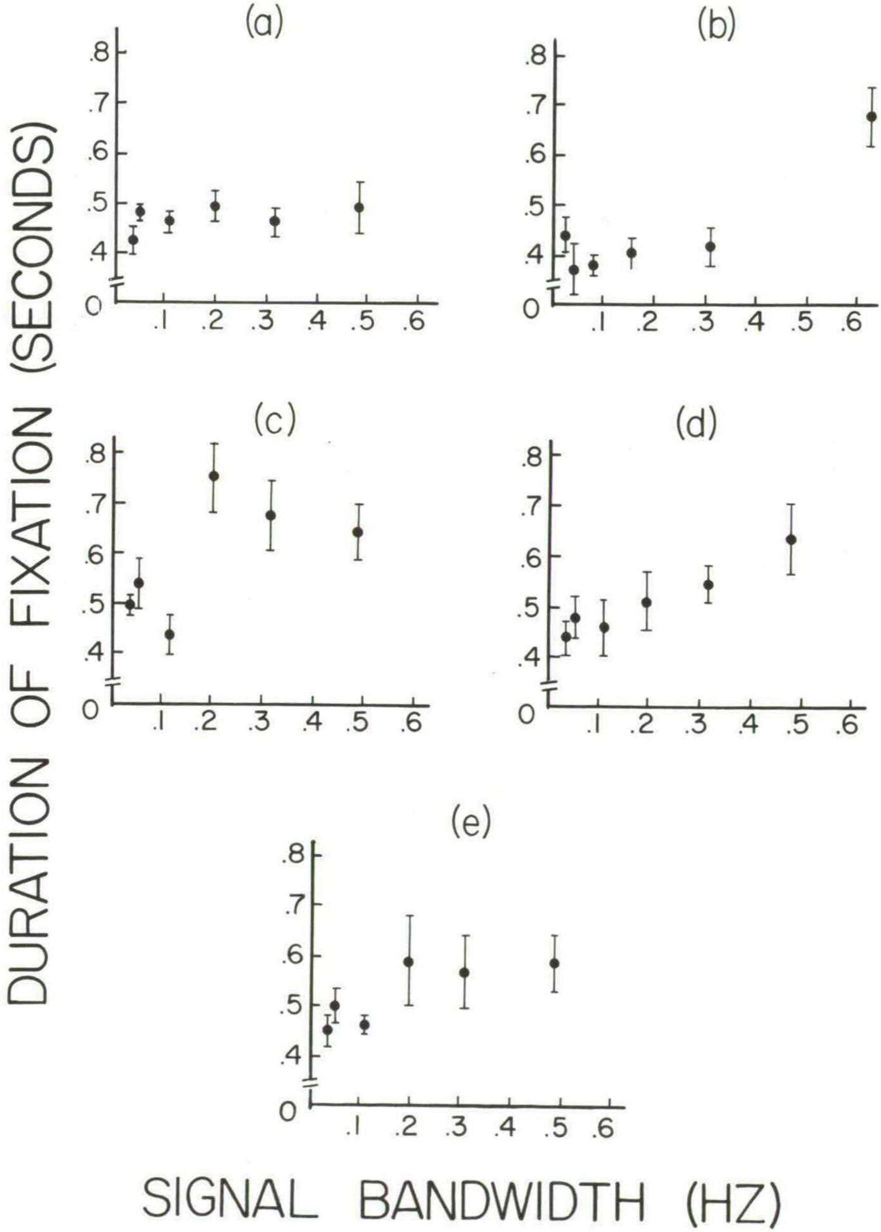

(a)
(b)
(c)
(d)
(e)
.8
.7
.6
.5
.4
0
.1 .2 .3 .4 .5 .6
DURATION OF FIXATION (SECONDS)
SIGNAL BANDWIDTH (HZ)

Figure 5.

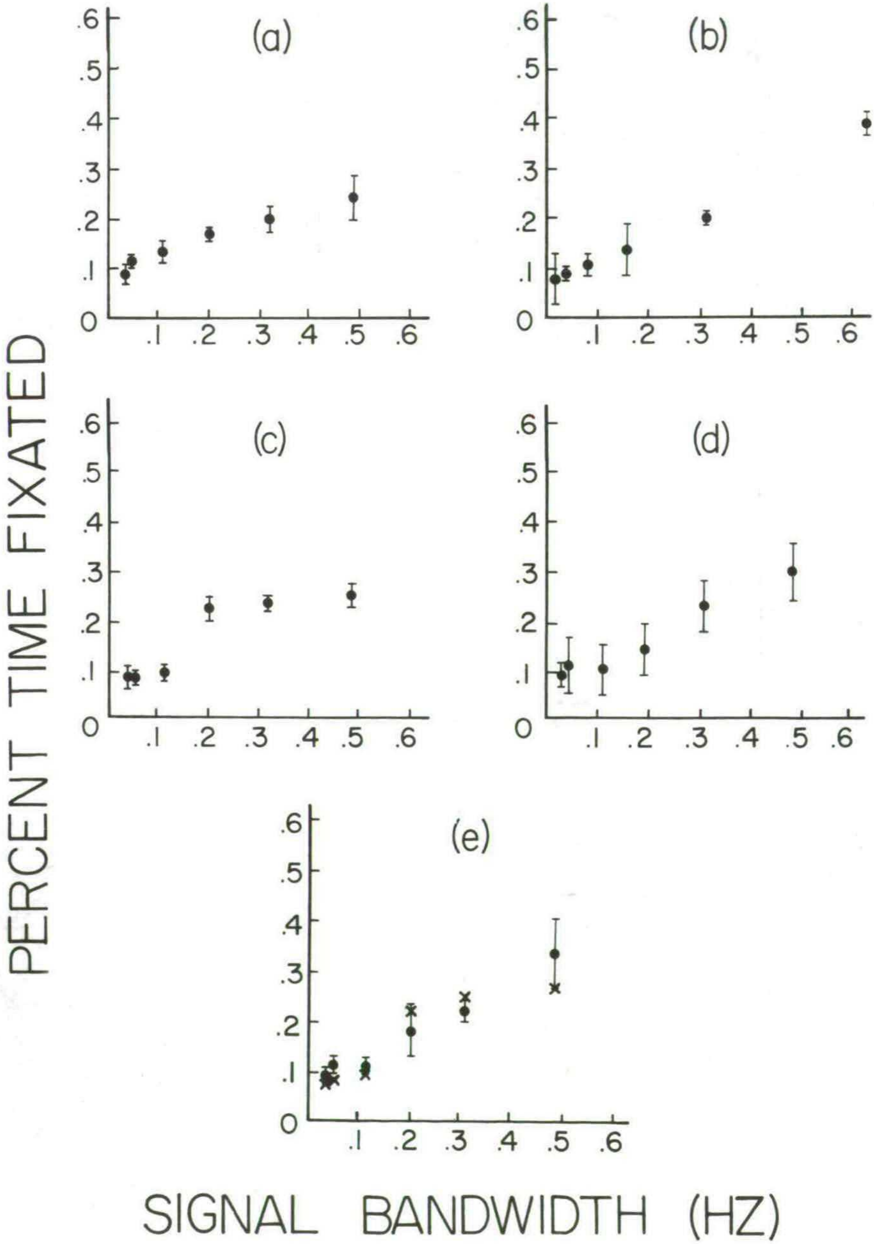


Figure 6.

CHAPTER IV

THE IMPLICATIONS OF SCANNING MODELS FOR THEORY AND APPLICATIONS

1. *Information Processing.*

The scanning theory itself and the models which derive from it are fundamentally theories and models of human information processing. The difference between these models and those previously dealt with in the behavioural sciences lies largely in the fact that the information sources are continuous in nature rather than representing discrete arrays of stimuli with varying probabilities. This difference in emphasis leads to a different mathematical structure and a different analysis of the role of the human operator in information processing. Since the scanning theory deals with large numbers of continuous information generators, the human observer is constrained to commutate amongst the various sources and to allocate time in some way that minimizes some cost functional; presumably at the end of a long period of training the operator, either consciously or unconsciously, has arrived at some allocation algorithm which allows HO to minimize the number of errors or to minimize the amount of uncertainty HO has with respect to any of the sources. For continuous sources, as indicated in Chapter II, it is possible to break down the general statement of the information generation rate of a continuous signal into two minor equations which predict the frequency and duration of visual samples taken by an ideal observer.

We would expect to find that with increasing information generation rate, the total amount of time spent on a source of the information would be proportionally larger and indeed, as indicated in the four six-dial experiments, the percentage of time spent is a nearly linear monotone increasing function of signal bandwidth. Since in all cases there was no specification of the required precision, we can only assume that subjects treated all signals identically and assumed a uniform precision requirement for all. Under these conditions we would expect a linear relation between observing time and signal bandwidth. The correlation overall is sufficiently high to indicate that this expectation is fulfilled. Typically, short saccades last about 40 ms. The intercept of the function, at approximately 7 % time, suggests that when there are distributed sources, about that much time is wasted merely moving the eyes from place to place. This is turn gives some indication of what might be gained by combining information sources into a single instrument.

From the very design of our experimental procedure (and the preliminary assessment of the ability of the observers to see any dial other than the one fixated upon) it is clear that the effective stimulus for the change of the point of regard is internally generated. The decision for the practiced observer is hardly in the nature of a voluntary one. That is to say the subjects in these experiments were not actively directing the eyes to move from one dial to another. Rather it is as if the eyes' mind, earlier hypothesized, directed the eyes in such a way as to bring to attention what the mind's eye wanted to see.

The theory is actually a good predictor of the behaviour of the subjects. In all cases there was, as there should be, a monotone increasing frequency of fixation with increasing bandwidth. All the models predicted this increase and all the data showed it. Yet in all the data there was a minimum fixation frequency which is predicted only by the Rational Sampling Model (RSM).

All the models except the RCM are based on the idea that the observer's uncertainty about the state of a system variable increases with increasing time from the last moment when the instrument showing that variable was fixated. Even the periodic sampling model is based, in a mathematical rather than a psychological way, on this requirement. The uncertainty is generated by the fact that the signal generates information all the time. When it is not being observed, this information is equal to the uncertainty created in the observer. This will be true only if the observer has a well quantified concept of the statistics of the signal. The fact that there is such a strong linear relationship between bandwidth and proportion of time spent indicates that the subjects have acquired such a concept. What is puzzling is that the rate of increase of time spent is as low as it is per unit increase in bandwidth.

Underlying the theory and the models is the tacit assumption that human beings are uncertainty reducing machines. Their behaviour the world is directed toward organizing and ordering the chaos that they find around them. Learning is an organizational process. In particular, the learning of the statistics of an array of signals permits progressively improved sampling tactics to be used within the overall uncertainty reduction strategy. The result of the improved tactics is a reduction in effort or an increase in information gained and a concomitant reduction in the uncertainty of the observer.

The instruments in the experiments were identical with respect to the required accuracy of reading. The only thing that differentiated one from the other was th: bandwidth of the signal that drove the pointer. In order to make a rational allocation of visual attention to the various signals, the observer must learn the bandwidths of those signals.

Since the subjects received as many as thirty hours of exposure to the array of instruments during the experiments, there was what would in normal circumstances be considered adequate time to learn the signal statistics. However, since the real life tasks to which we wish to extrapolate the results may involve thousands of hours of experience before competence is recognized, our results may be valid only in the laboratory environment. Psychological experiments as a rule do not use learning times which even approach those encountered in the working world. In the four-dial experiments the relative frequencies of fixation on the various dials are roughly in accord with the model after about three hours of exposure. However, although the sampling was at the "correct" frequencies, it was not effective for the detection of deviant readings on the instruments as indicated by a relatively low detection rate.

Thus, something was lacking in the performance at three hours. It is probably the case that something was still lacking after as many as thirty hours of experience.

What engenders the decision to look ? If the eye is fixated on dial i and moved to dial k , the question is : why did the observer look at this moment at this dial ? The models, except for the RCM, propose in various ways that the increasing uncertainty about each of the signals not looked at ultimately leads to a decision on the part of the eye's mind to look at the signal most "needful" of attention; that is to say, at the signal about which the uncertainty is greatest. Where does the uncertainty come from? There are two sources :

1. The signal is random and may vary during the period of non-observation over a range which increases with increasing time. The observer has some strategy, and the strategy determines the prediction of the model.

2. The observer forgets what was seen on the last observation so that uncertainty about the reference value of the signal will increase with increasing time.

Each of the models uses one or both of these presumed sources as a basis.

It is noteworthy that the data for percent time as a function of bandwidth are so nearly linear and almost identical for all experiments, both in terms of intercept and slope. The results of this measure, percent time spent, are more revealing with respect to the manipulations imposed on the .12 bandwidth signal in experiments 4 and 5. It will be recalled that virtually no alteration in sampling frequency was observed. However, the durations associated with the two manipulations were reduced. These manifest themselves in a reduction in the total proportion of time spent on the .12 signal in experiments 4 and 5.

2. *Scanning.*

Posner (2) and Moray (7) both suggest that scanning processes occur internal to the observer. Thus, Moray's work deals with the switching from one ear to the other when signals are simultaneously available at both. Posner's explores the movement of the mind's eye from one part of the visual field to another. In both cases the switch takes a finite amount of time. The analogue between these internal scannings and the overt visual scanning is clear.

It will take a significant research effort to explore the question of whether the internal scanning process follows the same rules as the external one. One might well imagine that it would on the grounds that the maximum efficiency of observation can be achieved by the utilization of an algorithm similar to the models presented. These works suggest that simultaneous attention to information sources probably is not possible even if both are present in the visual field at the same time. In our case, however, there is no question as to whether an observer can deal with many sources simultaneously : he can look at only one place at a time.

It is tempting to take the uncertainty based observing concept and apply it quite generally to all visual scanning, for example, the scanning of static, pictorial scenes. It is possible to assign all of the uncertainty which arises in the course of time to the failure of short-term pictorial memory. Thus, an initial fixation reduces uncertainty at the point of regard to some very low value, and to a lesser degree at points in the scene removed from the point of regard according to the decline in visual acuity with angular deviation from the line of regard. One could, therefore, convolve the visual acuity function with an information density function derived from a two-dimensional spatial Fourier analysis of the scene and arrive at an information density map available to the observer through the visual system. The selection of the next point of regard would depend upon the particular strategy or tactic chosen. If the tactic were merely to look at that part of the visual field which had the maximum uncertainty about it, a deterministic scanning model would be developed. In the steady state solution it will be assumed that all points in the scene have been either fixated or within the visual field during a fixation, so that the decay of information or increase of uncertainty with time since last observation can also be calculated. Under these conditions, the tactic, either deterministic or stochastic, will yield scanning behaviour.

3. *Probabilistic Versus Deterministic Scanning.*

The Periodic Sampling Model assumes essentially deterministic scanning. Once a sample has been made, the next sample would be made at a fixed interval and it would be the observer's task to organize the successive samples in such a way that the interweaving of the information would meet the requirements of the model. This is probably impossible to do except on very few, carefully constructed sets of signals. In all the other models, the tactic can be arranged to be either deterministic or stochastic and to allow the observer spare time or not. That the observer has spare time means that when the queue is empty, that is to say no instrument is waiting for attention, H.O. may do nothing at all. Still, the next item to appear in the queue is fully determined by everything that has happened before and by the nature of the signals themselves. If there is no spare time, it is assumed that whichever signal has the greatest uncertainty will appear in the queue. The observer always has something to do and some signal will be looked at even though its uncertainty may be relatively low. Thus any fixed threshold mechanism to trigger viewing behaviour will lead to deterministic sampling. The first signal to penetrate the threshold captures the attentional mechanism. In the other case, whichever signal has the greatest uncertainty at the conclusion of the current fixation captures the attentional mechanism. Alternative to this is the notion of the sampling model based upon random selection, given that the proportions in the hypothetical urn of markers with the names of the various signals, vary according to the instantaneous uncertainty associated with these signals. The observer pulls from the urn a marker which indicates which signal is next to observe. It is possible that a signal with very high uncertainty will not be observed on a particular occasion and that some instrument of lesser uncertainty will be observed. However, since the uncertainty will continue to increase the probability that the unobserved instrument will be observed grows with successive passings over. In this model, of course, there is no spare time unless "blank chips" appear in the urn. The questions then are : How did these chips get there and how are they caused to disappear as indeed they should when the perceptual scanning load is high ? One might imagine, for example, that the blank chips would be in proportion to the spare capacity. An attractive alternative solution to these questions is that behaviour with a low load situation is dependent upon the greatest uncertainty or no-spare-time deterministic mechanism and, as load increases, moves to a threshold uncertainty.

4. *Workload, Queueing and the Organization of Behaviour.*

Scanning obviously represents a significant load. The orginal development of the models presented was engendered by an interest in perceptual loading of aircraft pilots. When the load is low, whatever the mechanism of selection, there is no time pressure on the observer. As the load gets high, however, time pressure may arise and two or more signals may simultaneously call for attention. Under these conditions, the operator must allow a queue to form with two or more "customers" in the queue and deal with them according to one or another queue discipline. When the load is relatively low, but the queue still has two or more signals waiting service, then queue discipline might well be first-in-first-out (FIFO). If the load is very high, and the queue long, it might be tempting for the operator to choose a different queue discipline : last-in-first-out (LIFO). Finally, if the queue is very long and the load very high, breakdown of behaviour may occur and service-in-random-order (SIRO) may occur. These three forms of queue discipline may commonly be observed in the way one deals with different degrees of loading in paperwork as well as in monitoring. It seems reasonable to assume that the mathematics of queues dealing with these three forms of discipline would be useful as a basis of model and theory of human behaviour under overload conditions. The breakdown of behaviour into SIRO discipline would ordinarily be undesirable. The appropriate strategy might better be to adopt a system of priorities in which certain signals come, as it were, to the head of the queue and are serviced immediately. Here again, the relationship between queueing mathematics and psychological theory is clear and worthy of some attention. Finally, the dilemma, described in greater detail in the next section, dealing with the large disparity between the data obtained in these studies and actual practice in such enterprises as nuclear power plant control rooms, suggests that compensatory behaviour is practiced by the operators in these control rooms to reduce the scanning load. It is probable that some form of "chunking" occurs such that the operator uses knowledge of system dynamics to be selective about the things looked at. If the coupling between chunk outputs and system states is sufficiently high, it may suffice for the operator to look only at the outputs unless the value of the output departs from the nominal state. All in all, traditional behavioural theory has relatively little to say about monitoring behaviour in operational situations. That which can be conveniently studied in the laboratory is of necessity small, short and clean. The real world is large, long and dirty and it takes a new kind of experiment and theory to deal with that real world.

5. Applications.

Scanning behaviour is ubiquitous in the real world. Virtually all modern tasks involve the interaction of human operators with various levels of automated systems. Thus, the principal task of the operator is to observe the state variables of the system; to make decisions and then to undertake appropriate actions. All of the experiments demonstrated that the sampling of low frequency signals was well above the theory when any of the theories was based upon stimulus generated uncertainty. It was necessary, therefore, to construct an appropriate model involving an increase in uncertainty from internal sources: the operator forgets what was seen. Yet if this conclusion is correct, i.e., there is a minimum interval below which the monitor cannot go, then these data and the uncertainty threshold model which appear so rational for the restricted laboratory task cannot hold for the complex real situation. In the real world, a team of perhaps four observers may be confronted by an array of as many as two thousand indicators and signals. In order that they can cope with the task, the "floor" created, apparently by forgetting, must be absent even with uniform intervals between observations and with uniform distribution of the monitoring load among the four monitors. There are 500 instruments per monitor which must be dealt with. If a monitor can look 2 1/2 times per second (and this would be a steady 100 load) there would be a delay of 200 seconds from one observation of any indicator to the next observation of that same indicator. This is more than ten times the intervals permitted by our subjects with what was presumably a full load task. The extra looks at the low bandwidth dials were obtained at the expense of the higher bandwidths which were looked at less often than rational theory would demand. Yet the four operators usually appear to be underloaded with their two thousand indicators. How can one interpret this observation? The operators may in reality be overloaded and unaware of it. They may be overloaded and resigned to it, or they may in fact be underloaded by virtue of having restructured the task so as to make it "do-able". Could an operator be overloaded and unaware of it ? In a sufficiently reliable system, the exceedingly low failure probabilities provide only little experience with emergency requirements. The failure to observe many of the instruments frequently has not resulted in system failure or degradation. There is no cause for alarm in the loss of information from the low bandwidth signals that remain too long unobserved. Adaptation in the sense of an automatic and unconscious process may permit unawareness of overload by permitting a shift of performance criterion. This result would be , of course, a dangerous one. The operators are lulled into a belief that all is well when in reality all is not well. It is merely that not enough time has passed to permit the full scope of events to occur to most operators. This leads to what might well be an aphorism of complex systems that *the more reliable the plant the less reliable the human operator.*

Could an operator be overloaded and aware of it ? Although one can imagine this process, there has been no evidence of task induced breakdown amongst nuclear power plant control room operators and one might well imagine that operating under continual overload would lead to this. This alternative probably can be discarded. The most likely hypotheses are either that a different behavioural pattern is learned : system "chunking"; that memory is improved in some way; or that task design has been successful in altering the psychological requirements placed on the operators. The last is in most cases the introduction of arrays of warning lights which call the operators' attention to the deviations of the system and eliminate the need for memory of dial readings to serve an important function. Yet it is precisely these warning lights which elicit so many complaints from both control room operators and engineering psychologists. If memory were the loading process, which it seems to be in the small laboratory task, then the lights are necessary. On the other hand, if there has been developed a new pattern of behaviour which can overcome the problem of forgetting, the lights may indeed be unneeded. The coherence of the plant may make it possible for such reorganized behaviour to occur.

Whether the task be that of controlling a nuclear power plant or aircraft piloting, there are subsets of signals which may be monitored by observing only the output stage which, if it is within bounds, implies that all the previous stages are also satisfactory. For example, if the observation of altitude shows that altitude is within bounds, it will be unnecessary to monitor rate-of-climb as frequently as might be predicted purely on the basis of the statistics of this latter instrument. In the nuclear power plant, the maintenance of core temperature may make it unnecessary to examine control rod position or heat transfer with as much detail as would be the case if one were interested in only rod position or heat transfer by themselves and had no appreciation of their effects on other instruments In all probability, those laboratory studies which use randomized functions and attempt then to extrapolate to the highly constrained, serially organized behaviour of coherent systems, will yield excessively conservative results. In all probability it will be necessary, if one is to explore behaviour in nuclear power plant control rooms, to undertake the gathering of data either in a simulator or in a control room itself. The extrapolation from laboratory data to the operational situation is likely to be quite unwarranted.

Despite all the foregoing, it is clear that in a task with a relatively small number of indicators, the information generation rate of the various signals will determine the amount of time in aggregate spent on each and the bandwidth of the signal will determine how often each is looked at. From this latter set of data, the Link Values or transition probabilities can be calculated and optimum layouts computed. An example of this is the work by McRuer, Jex, et al (8) which demonstrated that even the simplest approximation of observation frequencies and transition probabilities sufficed to specify a layout of instruments on an aircraft panel which was in precise accord with actual practice. We may have faith, therefore, in the utility of these models even if their underlying theory appears to be deficient with regard to many of the complex tasks of real life. Ultimately, data from the operational situation will have to be gathered in order that improved models more appropriate to tasks like that of the control room can be created.

6. Research Problems.

The number of possible research questions which can be raised is unbounded. Many variables which were controlled to be constant or left uncontrolled in the experiments reported here, could be swept over a wide range and the effects measured. Alternative models for sampling behaviour crowd the mind as one thinks of the diversity of possible strategies. What is lacking is a data base. We have measured the behaviour of observers monitoring four dials and six dials. The results for even these similar circumstances are quite different although there are consistencies as well. For example, in both cases there is an oversampling (compared with most of the models' predictions) of the low bandwidth signals.

The data from the Pilot Eye Movement Studies deal with the behaviour of pilots monitoring larger numbers of instruments, but there are problems in attempting to fit models to them. The characteristics of the signals are unknown. Further, the pilots were controlling as well as monitoring. Both of these make a comparison of the results difficult. The lowest frequency of fixation in the flight situation is lower than that found in these experiments but this may be the consequence of system coherence or of control activity or of both.

The number of signals to be monitored makes a difference. Our finding of a "floor" frequency of about .1 observation per second is completely at odds with the manifest behaviour of persons performing everyday tasks. If that floor held for all situations, a nuclear power plant control room would require approximately 50 operators to maintain vigilance over the whole array of indicators. In reality, 3 or 4 people operate (apparently successfully for the most part) such plants. It may be the case that the operators are lucky, but this is not a credible assertion. There must be a logic to the many instrument watching task which can be modelled successfully. The principal difficulty arises from the fact that we don't know what reality is. We have no data on observing behaviour of highly skilled operators performing a task involving thousands of signals to be monitored. It may well be that the power plant operators conform to a conditional sampling model and have good memories for the state of the observed variables. It is therefore of the utmost importance to collect such data.

Another basic difficulty stems from the fact that laboratory subjects are not well trained in the industrial sense. Few experimenters are willing to commit a full year to the training of a reasonable number of subjects in the operation of a highly complex simulated plant, and then to begin what may be as long a period of data collection in order that suitable statistics can be computed; and few subjects are willing to work for a year or two learning and demonstrating a useless skill. The tendency is and has been to perform brief experiments and extrapolate the results to situations of far greater complexity and to subjects of far grater levels of skill than any laboratory subject can possibly achieve. In all probability, the only data which can serve our need are those gathered in the field. Field gathered data have their own problems of course, but at least they are derived from trials which correspond in important detail to the focus of application.

The models that have been presented here lie at various points along a dimension of "rationality". The Conditional Sampling Models have a high degree of face validity, whereas the models that assume that the H.O. is trying to reconstruct the signal do not, except for the assumed forgetting. The data say otherwise. The data, except for the lowest bandwidths, are below the Periodic Sampling Model, whereas all the predictions of the conditional models are to the left of the periodic, above it. A puzzling question arises when one considers the data on the relatively large bandwidths of .32, .48 and .64 hz. Here the sampling behaviour is well below the line with a slope of 2. The explanation offered by the Rational Sampling Model is that velocity is perceived and that although there may be an increase in duration because of the necessity for extracting the additional information, there is simultaneously a decrease in frequency of observation from that which would be required if only the value of the variable were perceived. Examination of the actual data shows that for the most part the high frequency signals, .48 and .64 hz , do have a longer duration of observation. Further, the percent-time-spent data strongly suggest that the process is, in fact, linear with regard to information processing rate. Once again, we find that information Theory provides an adequate explanation of the data obtained despite its many shortcomings as a general theory of human performance. Experiments need to be devised, no mean feat, which will enable us to test matters relating to precision of readout without confounding the issue with ordinary boundary perceptions, such as are implied by the conditional sampling models. It may be then that we will be able to discriminate between the periodic or data-base sampling models and the conditional or state-variable sampling models.

CHAPTER V

CONCLUSIONS

Taken in the aggregate the results of the six-dial experiments, as well as those of the four-dial experiment, give very strong support to the general theory that the frequency with which instruments will be observed is largely determined by the statistical characteristics of the signal. In particular, frequency of observation is a monotonic increasing function of signal bandwidth. Because the signal powers are equalized, the correlation between signal bandwidth and frequency of fixation will be very nearly 1.0 . In all the experiments, the frequencies with which the low bandwidth signals were sampled were higher than those predicted by any of the models. As a consequence, the regression of fixation frequency on signal bandwidth has a slope in all the experiments in the vicinity of .6. The four slopes are very nearly identical, the least being .56 and the greatest .64. Similarly, the intercepts are in the small range between .15 and .22. The correlations are all high between sampling fixation and signal bandwidth, ranging from a least value of .959 to a greatest value of .989. All four experiments yield virtually identical results insofar as frequency of fixation is concerned. This speaks well for the consistency of the subjects and for the robustness of the data of the individual experiments. Even more striking are the data with respect to percentage of time allocated to the various signals. Here we find uniformly high correlations, the least value being .90 for Experiment 4 and the values for the other three experiments being .97, .99 and .99. The intercepts range from .06 to .09. Once again, there is great consistency. Thus, for both the four- and six-dial experiments the general findings are in substantial agreement for the two numbers of dials presented and for the increased range of signal bandwidth impressed upon the dials.

The four studies were simple replications with only minor changes in the way in which information was presented. For the most part, these changes had little discernible effect upon the ways in which the subjects distributed their attention amongst the various signals. In particular, the experiment which attempted to explore the effect of correlation between two of the signals produced relatively inconsequential results, However, it must be emphasized that in real systems relationships between signals reflect a physical relationship between the variables presented. Thus, in an aircraft we would expect that the correlation between altitude and rate-of-climb stems from the exact physical relationship such that altitude is the integral of rate-of-climb. Similarly, we would expect that the rate-of-turn indicator will have a strong relationship to its integral : heading. The magnetic compass and the gyro compass will be closely coupled and reflect almost the identical signal. The dichotomization of a continuous signal did produce changes in the observing behaviour of the subjects. However these changes were not consistent. For one subject, there was a marked increase in the amount of attention paid to the dichotomized signal. For three of the other four there was a marked reduction. In the aggregate, there was a slight reduction in the amount of visual fixation paid to the dichotomized signal. The magnitude of the change however, was not large enough to make any important change in the calculations that one might wish to make from the general model to a real system.

Overall, one might take as a conclusion that it is the nature of the signal which drives the display rather than the method of display itself, and that the bandwidth is the signal most important factor influencing the frequency of fixation. This will be true particularly if the powers of the various signals which are being sampled are identical. Naturally, since this is predicted by all the models derived from the general theory, it is not too surprising and it is intuitively satisfying that the results evolve in this way. What is less than satisfying is that the regression of fixation frequency on bandwidth has such a low slope. It will be recalled that periodic sampling at the Nyquist interval should yield a slope of 2 and that all of the conditional sampling models yield slopes higher than 2. Finally, the random costrained model will yield a slope of less than 2 only under exceptional circumstances. In a sense, then, all of these models are falsified; none makes an adequate prediction. It may well be the case that the slope is reduced by virtue of the oversampling of the low frequency signals, but it is also the case that all the high frequency signals are undersampled insofar as the various models would suggest. The only model, therefore, which satisfies the data and is in general accord with the theoretical notion, is the model which assumes that sampling takes place whenever a threshold of uncertainty is reached and further that this uncertainty may develop either on the basis of forgetting or on the basis of signal characteristics, per se.

Yet even this attractive model, with its equally attractive adherence to the data obtained in the present study must fail for tasks involving very large numbers of instruments. The uncertainty induced by forgetting would lead to staffing requirements greatly in excess of actual practice in a wide variety of industrial situations. This might arise because people are not subject to the uncertainty caused by forgetting or it may be that in fact personnel in these complex machines do not sample all the instruments but only a very few. In that case the kind of model proposed might fit. It would take the acquisition of operational data from a suitable task to make possible the selection of one of these models or the creation of a new one. In the absence of a data base, little can be said about this situation, other than the fact that the behaviour in the control room cannot simultaneously be of the form found in these experiments and deal with all the instruments. Either the behaviour corresponds more nearly to the conditional models, or many of the instruments will be ignored.

If the Rational Sampling Model (RSM) is to be accepted as the underlying basis, then it would appear that for these dials, at any rate, the transition from the forgetting mode to the velocity perception mode must occur with little space between, if any at all, for the Nyquist sampling mode in which only the value at the moment of sampling is detected by the observer. The data of Experiments 4 and 5, in which minor manipulations of .12 hz. signal were attempted, are more in accord with the RSM in that the two or three points for the lowest bandwidths are essentially horizontal and the slope for the two highest bandwidth signals is very much lower than for the intermediates, so that a three-part function could be an adequate fit for the data of these two studies. On the other hand, the data for Experiments 2 and 3 suggest a two-part model as being a better fit.

The data on durations are not conclusive. Without a precise knowledge of the amount of uncertainty associated with each of the signals, it is difficult to make exact calculations as to what the duration of observation should be given some constant processing rate. In particular, since the bandwidth is known, the unknown part, the relative accuracy or the ratio A/E, would have to vary for the various signals in order to make the duration figures conform to the theory or to make any sense at all. Thus, for Experiment 2, the durations are very nearly constant for all bandwidths whereas for Experiment 3 they are a rising function of frequency and are constant only when transformed in accord with the random constrained model. This is also more or less true for Experiments 4 and 5 so that no general conclusion can be reached. When all four studies are combined, we get data which deviate markedly from what we would expect on theoretical grounds, irrespective of whether we use the raw data or the random constrained data. The data on percent-time-spent as a function of signal bandwidth are consistent both for the tree studies which use the .03 to.48 signal bandwidths and the study using the .02 to .64 signal bandwidths. The data for the latter study fit neatly into the function for the data of the other three studies. In general the percent-time functions suggest that whatever deviations may have occurred with respect to frequency and duration of observation were neatly compensated such that the percentage of time spent of the signals is very nearly lineary related to the signal bandwidth. The correlations are all very high and the points cluster closely around the lines, with the exception of the data for Experiment 4, where the residual error reaches the value of . 033. The intercepts are all in the same general area and suggest that about seven percent of the available time is spent doing something other than looking at instruments, presumably in moving the eyes from one place to another. Figure 6e shows that the data for Experiment 3 involving both the lower and the higher signal bandwidths correspond well to those for the aggregate of the other three. Thus, we can say with considerable confidence that the amount of visual attention paid to a signal in the midst of other signals is proportional to the signal bandwidth and therefore to the information generation rate. These results provide, once again, strong support for the idea that Information Theory is a valuable metric of at least some kinds of human behaviour. In our case, it provides a perfectly sound base for estimating the percentage of time that an observer will spend monitoring a signal.

It is unfortunate that the values of the signals at the moments of observation were not recorded. If they had been, it might have been possible to evolve an alternative strategy and a new conditional sampling model which would provide a better basis for the data obtained. The conditional sampling models have strong intuitive appeal to them. It would seem most reasonable that a highly trained monitor of many signals would take advantage of the fact that knowledge of what had been seen would be useful in determining when to look again at that signal. This is particularly true in light of the fact that the observers were instructed to signal when the pointer reached or exceeded the limit. In other words, the conditions of the experiment were more nearly those appropriate to the conditional sampling models, and we would have expected that those models would be a good description of behaviour. We abandon them with reluctance. The rational model also has a high intuitive appeal, but it lacks the notion that the observer does anything with or about the value of the signal observed except in the course of some ten or fifteen seconds to forget what it was.

These studies have, if they have done nothing else, pointed out the need for more extensive research in the signal monitoring behaviour of human beings. Clearly, if we are to understand *why* people do what they do, we must learn *what* they in fact do. It is necessary to record not only the positions of the eyes but also the value of the signals which are observed. It is only the relationship of these two sets of data that will tell us whether there is anything at all in the idea that observers make use of the information that they see in deciding when to look again. Another experiment which would be most useful would be one that used the eye movement to lock the instrument pointer in its position for the duration of the fixation. This would eliminate any possibility that rate information could be gathered. Under these conditions, we might expect significant modification of the functional relationship between frequency of observation and signal bandwidth. In particular, if the "Rational" model holds, and it appears to be a better description of the data than any of the others, we should be able to manipulate the decreasing slope of the high frequency signals through the use of this strategy. However, it must be recalled that although the frequency data tend to curve concave downward, the time data, by and large, are straight, suggesting that there is a compensation for the decreased frequency of viewing in the form of increased durations. The use of the "sample and hold" mechanism, might solve this problem.

There needs also to be a serious effort to collect eye movement and eye fixation data in complex plants like nuclear power plants and chemical process plants. It is absolutely certain that the kinds of findings which we have observed in these studies *cannot* be representative of the viewing behaviour of the monitors of these plants unless a totally different strategy has been employed involving the monitoring of only a few critical instruments leaving the large majority unattended to. If this proved to be the case, it would have a marked effect on the design of control panels for human operation. On the other hand, if the data demonstrated that all instruments are looked at, then it is clearly the case that some instruments, perhaps most, must be looked at only very infrequently indeed and the data may conform more closely to one of the conditional model or to the random sampling model. Certainly they cannot conform to the predictions of the rational model.

Lastly, experiments have to be done on physically coherent systems. Although it is customary in experimental psychology to use randomized events independent insofar as possible one from the other, the real world is composed of systems possessing physical coherence which obey physical law. Systems can be incorporated into the mind of the observer in the form of the mental model which guides the monitoring behaviour. In the case of independent random signals no mental model can be constructed . It is entirely possible that the results obtained are quite specific to circumstances in which no mental model can be constructed howevere long the period of exposure and in which the signals are statistically completely independent. The validity of the general theory is established. The construction of particular useful models must await the development of data bases from coherent systems involving larger numbers of instruments as well as the numbers used in these experiments.

CHAPTER VI

LOOKING BACKWARD

My interest in visual sampling grew from a dissatisfaction with the work, current at the time, on the Human Operator as a controller. Some four or more years of experimentation and analysis had seen the complexity of the case studied move from one-degree-of-freedom systems to two-degree-of-freedom systems. Two axis tracking was almost beyond the capability of the models and modellers. Yet it was clear that aircraft instrument panels were continuing to increase in complexity and that the role of the pilot was changing from controller to monitor of automatic systems.

Information theory was tnen very much alive but limited, in its application to psychological problems, to consideration of discrete signals arranged in space or time, ot both.

The Pilot Eye Movement Studies were being carried out but in a quite non-analytic way. That lack of analyticity was displeasing. When, upon re-reading Shannon (3) I saw again the appendix on sampling, I experienced, in the context of the Psychology Branch of the Aero Medical Laboratory at Wright-Patterson Air Force Base, a sudden awareness of the significance of the sampling notion for the understanding of the visual scanning behaviour of human beings. The formulation of the information generation of continuous signals plus the sampling theorem meshed with my interest in the problem of instrument scanning. It seemed obvious (even if later to be proved not completely correct) that if mathematical law dictated the number of samples required for the sampler to know the signal, it applied as well to humen beings as to abstract sampling machines.

Henry Quastler's conference on Information Theory in Psychology provided a convenient motivation for the preparatior of a short paper which appeared in the proceedings of the meeting in a book of the same title. (Reference 14 of Appendix A). To the best of my knowledge, only two people read it in the seven years following the publication. The paper was aimed at a number of targets : instrument panel arrangement, pilot workload, in vivo assessment of instrument design, in vivo assessment of pilot scanning strategy. If one could arrive at analytical, ideal behaviour toward an instrument on the basis of the characteristics of the signal which drove the instrument, then departure from the ideal could be taken as a direct measure of the quality of design or of the behaviour. It would obviate repeated measurement or eye movements in new aircraft, new maneuvers, new cockpit designs. Importance, rather than being a measure derived from behaviour, would be determined by system analysis of what needed to be done and when and to what precision.

The workload problem had been talked about but mostly in an anecdotal way. It was common to see in the press, and also to hear in informal discussion at technical meetings, concern expressed about the ever-increasing number of displays in the cockpit of aircraft and the perceptual difficulties the pilot must be having. At the same time the pilots seemed to be coping very well with the situation. The apparent high perceptual load did not appear to be real. If one could calculate the loads imposed by the signals entering the cockpit and add them up, one might be able to arrive at a quantitative instead of an impressionistic assessment of the load imposed on HO by a system.

I confess that I had hoped that other people would read the paper and be motivated to do the confirming experiments but in the end I had to do them myself. The result was the Four Dial experiment. It was not, in retrospect, well designed. However, at the time (about 1955) technology was not what it is today (1983). There were no on-line data processors, miniature bio-electrodes, solid state amplifiers, digital recorders. I felt that the usual extrapolation from subjects with thirty minutes of training to pilots with 3000 hours of experience was risky business. So I trained my subjects for more than 30 hours and took data along the way in order to find out how long it took for them to stabilize in their scanning behaviour. Indeed it took about ten hours for scanning to stabilize and more nearly twenty-five for detection to arrive at a reasonable high level. The extrapolation from thirty hours of subject time to the 3000 hour pilot was risky enough but more time was impractical.

The volume of data that had to be processed was prodigious. This was especially true when one considers that everything except the final statistics was done by hand. The photographic recording method was reliable (in terms of inter-reader agreement) but time-consuming. The films had to be sent out for processing and were returned a week or ten days later. I never really knew what had gone on or wheter any data had in fact been recorded even though I was planning the next recording sessions. It was necessary to stay nervous, uncertain and hopeful throughout the experiment.

The results were almost too good. The fine agreement between prediction and data made me feel that nothing more needed to be done and there the matter sat for about seven or eight years.

In 1962 Dr. Rayford T. Saucer of N.A.S.A. - Langley read the published account of the Four Dial experiment and proposed that the value of the model could be so great that it would be worthwhile to repeat the experiment with some changes and confirming the results. N.A.S.A. funded the study with what seemed like a vast sum of money and the Six Dial experiments were planned and run as described in Chapter III.

Even by 1962 no other work on the scanning problem had come to my attention. All we had as a basis was my own earlier results. I offered the problem of conditional sampling to M. Grignetti and J. Elkind as a minor challenge and they produced the analysis presented in the earlier publication of the project, NASA CR-434 in 1964. R. Smallwood and J. Carbonel also undertook to work on other parts of the general scanning problem. The former attacked the "mental model" of the monitor; the latter did a substantial analysis of the queueing notions that I had proposed. Since 1964 a number of investigators have studied, both analytically and empirically, the scanning problem. Among these were T. Kvalseth, J. Carbonel and R. Curry. Since these latter, and others, followed my own work. I felt it inappropriate in a dissertation to report extensively on them. In any event, my colleague, N. Moray, is preparing a comprehensive review of models of the human monitor including my own. I recommend it as a key to the literature on the scanning problem (9). In hindsight everything could have been done better. The model development and exposition, the experiments, the data analysis all could have and should have been better than they were. Where possible, I have improved on what was done twenty and thirty years ago. Some of the needed statistics are not available for proper hypothesis testing beyond the original limited goals of the contractual work. Although the data exist in the form of punched cards (the Four Dial) and digital tapes (the Six Dial), it would probably be more economical to start all over again and, in the light of my own and others' work, to do everything correctly.

REFERENCES

1. Milton, J.L., Jones, R.E., and Fitts, P.M. Eye Fixations of Aircraft Pilots : Frequency, duration and sequence of fixations when flying selected maneuvers during instrument and visual flight conditions, AF Tech. Report 6018 (ATI 84010), August 1950.
   Note : this is one of a long series of reports on a variety of maneuvers and flight conditions. Still later reports followed on various aircrafts as well.

2. Posner, M.I., Snyder, C.R.R., and Davidson, B.J. Attention and the Detection of Signals, Journal of Experimental Psychology, General, 1980, 109, 160-174

3. Shannon, C.E. and Weaver, W. , The Mathematical Theory of Communication, University of Illinois Press, Urbana, IL, 1949

4. Senders, J.W., A Re-Analysis of the Pilot Eye-Movement Data, IEEE Transactions on Human Factors in Electronics, 1966, HFE-7, 2, 103-106

5. Fogel, L.J., A Note on the Sampling Theorem, IRE Transactions on Information Theory, 1955, IT-12, pp 47-48

6. Elkind, Characteristics of the Human Operator in a Simple Manual Control System, IRE Transactions on Automatic Control, May 1959, AC-4, pp 49-55

7. Moray, N.P., Personal Communication.

8. McRuer, D.T., Jex, H.R., Clement, W.F., and Graham, F.D. Development of a Systems Analysis Theory of Manual Control Displays, Systems Technology Inc, Hawthorne, CA, 1967 TR-163-1

9. Moray, N.P., Monitoring Behaviour and Supervisory Control, in Handbook of Perception and Performance, J. Beatty, ed., John Wiley, New York, to appear Fall 1983.

APPENDIX B

The Attentional Demand of Automobile Driving

ABSTRACT

A theoretical analysis and an experimental investigation of certain aspects of automobile driver information processing were undertaken. The theoretical analysis was the result of an effort to avoid difficulties associated with a servomechanistic approach to the automobile driving problem. The analysis is predicated on the assumption that a driver's attention is, in general, not continuously but only intermittently directed to the road. Between observation, uncertainty about both the position of his own vehicle on the road and the possible presence of other vehicles or obstacles increases until it exceeds a threshold. At that moment in time, the driver looks again at the road. This simple model appears to be a useful analog of the driving process. The analysis makes specific predictions about the form of the functional relationship between intervals between observations and vehicle speed.

The experimental program had two goals. One was the empirical investigation of the relation between amount of interruption of vision and driving speed. The other was the determination for various drivers and various roads of the values of some of the parameters in the mathematical model. This report presents the results of the theoretical and experimental investigation. In general, the model is a fair approximation of actual behavior and it remains for future work to determine whether this approximation is good enough to be useful for the specification of vehicle, highway, and user characteristics.

INTRODUCTION

All of us who drive are aware that at times we do not seem to pay very much attention to what we are doing. Drivers tune radios and light cigarettes; they blink and talk to passengers; they listen to the news or music. It is said that drivers become "road hypnotized" -- staring without seeing at the scene ahead. They look into rearview mirrors and scan for traffic police; they read advertising signs and search for turnoffs. Some of these activities are legitimate parts of the driving task; most are not. All of them constitute a diversion of attention away from the primary task of controlling a vehicle along a highway in accord with law and custom.

Driving is, in one sense, an error-free performance. No normal driver deliberately undertakes to get into an accident or into a collision. Collisions are (with some exceptions stemming from psychopathological origins) involuntary and accidental. For most driving situations, then, we can say that the driver accomplishes about what he intends to accomplish. The problems of analyzing error-free systems are great. In particular, where the system is a road, an automobile, and its driver -- for whom a preview of the path ahead is directly available -- it is difficult to use the techniques of simple linear servo-analysis. The output of the system is easily measured and easily understood, but it is extremely difficult to specify what the input is which results in the observed output.

Prior attempts have been made by other investigators to determine those elements in the complex visual world of road and traffic which elicit the driver's responses. Gordon(*1*) used the technique of restricting the driver's view in order to enforce head movements. Then a head-mounted camera photographed what the driver fixated upon, permitting the investigator to identify the important and salient features of the visual environment. Sheridan et al (*2*), rather than limiting the field of view bij artificial "tunnel vision", limited the forward field of view in two other ways. One way was to present everything out to a distance, d ; the other was to present only the segment of road from d to $d + \Delta$. The distance ahead that the driver can see, d , was controlled as the chief experimental variable, and his performance measured.

Our approach is to treat the driver as an information processing device. He takes in information visually through the windshield by observation of the entire road ahead, and transmits information by manipulations of the steering wheel, brakes, etc. Using this as a conceptual model, we have indulged in some speculation about the information processing task and have erected a not too complicated mathematical theory of how information flows into the driver and is processed.

Some of our theoretical notions arose from some personal observations made by the senoir author while driving on a straight road with little oncoming traffic. A heavy rain resulted in the windshield wipers being able to clear only a small sector of about 20 degrees behind the blade, so that visual conditions for the driver were somewhat analogous to those which would be presented by a radar sweep. The wiping speed was independent of the speed of the car. The driver became aware of a "psychological speed limit". Up to that speed, there was no anxiety; above that speed the driver became anxious and had to slow down.

This observation suggested a parallel between the sampling of a time function (in which the minimum sampling rate for signal reconstruction is related to the bandwidth of the signal) and the sampling of a road (a space function), where the minimum sampling rate is related to the characteristics of the road and to the velocity at which it is traversed. One might imagine that the road had a certain information rate built into it -- that is, there were so many bits per mile. The faster one traverses a portion of the road, the more bits per unit of time must be processed. Were the driver to see a road only at fixed intervals, he would develop uncertainty about what might have appeared on the road since his last observation, and about where he is on the road. If the intervals between observations were very long, then the accumulated uncertainty and the amount of information to be absorbed on the next observation would be greater. If the observation time itself were very short, the driver would be unable to completely reduce his uncertainty by absorbing the required amount of information.

The analogy leads to a parallel between information flow rate and some "equivalent bandwidth" of the road. The equivalence, of course, would be established bij the driver's selection of a speed at which he traverses the road. Given a fixed sampling rate, then, if a driver were to traverse a road at his "maximum" speed, one might argue backwards that the speed/sampling-rate combination would permit estimation of the equivalent bandwidth of the road and of the information density of the road.

There are a number of factors which can influence the selection of a driving speed. One would be the wiggliness of the road -- that is, the frequency with which the road deviates from a straight path by enough to require corrective action. Another factor would be the overall density of significant obstacles in the road. Still another might be related to how accurately the vehicle can be steered and the degree to which, without attention, it maintains the desired path. Less easily quantified is the driver's estimate of the probability that some new object will enter the road or the probability that opposing traffic will deviate into the path of his vehicle between observations. In general, all these factors can be reduced to an uncertainty estimate per unit length of road.

We have considered two experimental situations. In one the sampling rate is fixed and the driver modifies his speed according to the road, traffic, etc. The alternate is to cause the vehicle to travel at a constant speed and allow the driver to control the interval between observations at will. Thus, if the vehicle is very stable it does not need to be attended to as often as if it were very unstable; and if the uncertainty of steering is small, the driver does not have to look as often as he would if it were large. Similarly, objects at a distance produce less uncertainty than objects close at hand. One would expect that as opposing traffic approaches, the driver must attend more often. If there are many side streets, driveways, and the like, then the probability of cross traffic is high and the driver has to pay more attention.

Naturally, the rate at which one would look would not ordinarily be constant. Instead, it would be a continuous variable function of the instantaneous state of affairs. For example, if one were traversing a road at a constant speed and entered a populated area so that the probability of animal or human entry was high, then the frequency with which one looked at the road would go up. The more often the road turns, the more often must the vehicle be controlled, and the more often must the driver look in order to control. The point at issue is that a road demands attention. The attentional demand of a road is a characteristic of that road and of the traffic situation which may exist upon it as wel as the velocity at which it is traversed.

A rather important notion which underlies the theoretical work is that drivers tend to drive to a limit. We suggest that the limit is determined by that point when the driver's information processing capacity, either real or imagined, is matched by the information generation rate of the road, either real or estimated. The drivers may be wrong in their estimates, but they will tend to achieve this balance of input information rate and information processing rate. A driver in unfamiliar territory sees a great deal more uncertainty in the situation than a driver familiar with the territory. With familiarity there comes reduction of uncertainty, a reduction of information flow rate, and higher permissible velocity, granted the same territory and circumstances. This is reflected in the different ways people behave in automobiles in familiar and in unfamiliar terrains. It might be said that a curvy familiar road is "perceptually straight" since uncertainty about the road ahead is low.

Finally, drivers will accept different levels of risk and drive to a limit such that the probability of an accident is not greater than, but approaches, some upper threshold. Subjective acceptable risk level is a measurable characteristic of drivers and directly influences their behaviour on the road.

We have identified a number of factors which tend to control the speed of the driver traversing a road in the presence of traffic and other dynamic obstacles. These are, in brief : the width of the road and the frequency with which it turns; the estimated probability of intrusion from other vehicles and animals; the uncertainty associated with the vehicle dynamics; the precision of the steering mechanism; the residual errors of vehicle aiming; and a risk acceptance level which is a characteristic of each driver. We are then led to the development of a theoretical model of driver behaviour which describes and quantifies the cumulative uncertainty of the driver between looks at the road. The experimental program examined the actual behaviour of drivers with intermittently occluded views of the road.

AN UNCERTAINTY MODEL OF THE DRIVING SITUATION

The following derivation is based on steady-state driving with intermittently occluded vision. That is, the driver is assumed to have adjusted the vehicle velocity (or the period of occlusion) to meet his criteria of performance and risk. Thus both velocity and period of occlusion may be treated as constants.

Driver information.

Let the information density of the road be $H(x)$ bits per mile, where x is the road distance in miles. The reference point for distances $(x=0)$ is taken as the location of the driver at some reference time $(t=0)$. The amount of road information available to the driver depends on the limits of visibility and on the way in which the driver weighs the information. It is reasonable to assume that the weighing is a monotonically decreasing function of distance since the driver will attach more importance to the nearer sections of road (which require an immediate response) than to the sections further on. Let us assume, therefore, that the weighing function is $e^{-x/D}$,where D is the weighing constant in miles. This function meets the requirement of being monotonically decreasing and is mathematically convenient.

The driver, therefore, constructs a road information model that has an information density of $H(x)e^{-x/D}$ bits/mile. The maximum amount of information that can be stored by the driver is the integral of this density function over the range of visibility. The lower limit of useful vision, x_1 , is the distance one reaction time ahead, or the distance occluded from view by the vehicle, whichever is greater. The upper limit, x_2 , may be determined by external conditions or by limitations on the viewing time.

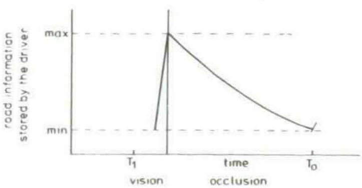


Figure 1. Timing of events

For ease of calculation, let the road information density be constant (H) over the section spanned by x_1 and x_2 . If we assume further that negligible error is introduced by taking the limits of visibility between 0 and infinity, the stored information reduces to

$$I_r = H \cdot D \tag{1}$$

As the vehicle proceeds over sections of the road stored in the driver's image, information containes in these sections becomes obsolete. At the vehicle velocity of V miles/sec., information becomes obsolete at a rate of HV bits/sec. Let us consider the situation in which vision is periodically occluded. The timing of events for one cycle is illustrated in Figure 1. At $t = -T_2$ vision is unobstructed and the driver absorbs information. The maximum amount of information is in store at $t=0$. Vision is obstructed for the folliwing T_d sec., during which time the store of information continually diminishes. The minimum amount of information is in store at $t = T_d$ at which time vision is restored and the cycle repeats.

During the period of vision the driver absorbs information at a rate of R bits/sec. Let us assume that R is constant as long as there is information to be absorbed. If the driver is able to absorb all the road information available before the period of vision has terminated, the rate of information input then drops to a level that exactly balances the rate of obsolescence.

There are two sources of information loss during the period of vision. (a) the rate of information obsolescence, HV, and (b) the rate of forgetting, which we shall assume to be proportional to the amount of information stored. Let the rate of forgetting be $I_r(t) / F$ bits/sec., where F is the time constant in seconds. If the maximum amount of road information available to the driver is HD, as derived in Eq. 1, the net instantaneous rate of information absorption during the period of vision is.

$$\frac{d}{dt} I_r(t) = \begin{matrix} R - HV - I_r(t)/F & I_r(t) < HD \\ 0 & I_r(t) = HD \end{matrix} \tag{2}$$

It will be assumed in the following dicussions that the driver is able to absorb all the information available by the end of the viewing period; that is, $I_r(t) = HD$ at $t = 0$

We shall now determine the amount of information stored by the driver at time t during the period of occlusion. The two sources of information loss remain, whereas there is no information input. The rate of forgetting is $I_r(t) / F$, as assumed above. Because of the way in which information is weighed by the driver, the rate of obsolescence varies with distance (and hence with time). Clearly, the rate at which information is "driven out" is related to the rate of forgetting. The faster the driver forgets, the less information there remains to be lost through forgetting.

The relationship between the rate of information obsolescence and the amount of information remaining in storage is

$$\frac{d}{dt} I_r(t) = -(V/D) I_r(t) \tag{3}$$

The total rate of information loss is

$$\frac{d}{dt} I_r(t) = -\frac{V}{D} I_r(t) - \frac{I}{F} I_r(t) = -((V/D) + (1/F)) I_r(t) \tag{4}$$

Given that the information in storage is HD at $t=0$, Eq. 4 yields

$$I_r(t) = H \cdot D \ e^{-((V/D) + (1/F))t} \quad \text{when} \quad 0 \leq t \leq T_d \tag{5}$$

Before proceeding with the development of the model, let us review the assumptions and derivations that have been made so far .

1. The driving situation is in the steady state. The vehicle proceeds at a constant velocity, V miles/sec.; the timing of looks is periodic such that vision is allowed for T_2 sec. and occluded to T_d sec.

2. The road has a constant information density of H bits/mile.

3. The information density of the driver's stored image is $He^{-x/D}$ bits mile.

4. The period of view is sufficient for the driver to absorb all the information available. The amount of information stored at $t=0$ is HD bits.

5. Information is forgotten at the rate of $I_r(t) / F$ and becomes obsolete at the rate of $I_r(t) V / D$ bits/sec. The amount of information in storage t sec. after the onset of occlusion is $HD\ e^{-(V/D+1/F)t}$ bits.

Driver Uncertainty

The driver is assumed to adjust the vehicle velocity (or the occlusion time) so that his uncertainty $U(t)$ is never greater than some criterion level U_c measured in bits. Since the uncertainty is greatest at the end of the occlusion interval $(t=T_d)$,the criterion may be stated mathematically as

$$U(T_d) \leq U_c \tag{6}$$

The driver uncertainty at time t sec. after occlusion is

$$U(t) = U_r(t) + U_n(t) \tag{7}$$

where $U_r(t)$ is uncertainty about the road due to a loss of relevant information and $U_n(t)$ is uncertainty about the position of the vehicle arising from random disturbances in the orientation of the car. The latter term takes account of the vehicle dynamics, the ability of the driver, and external disturbances such as wind and variations in the road surface.

The uncertainty about the road, $U_r(t)$ isequal to the amount of information that has been lost. It is equal to the maximum amount of information originally available minus the amount of information in store at time t. Thus

$$U_r(t) = H \cdot D\ (1-e^{-(V/D+1/F)t}) \tag{8}$$

This component of uncertainty starts from zero at $t=0$ and rises asymptotically to $H \cdot D$

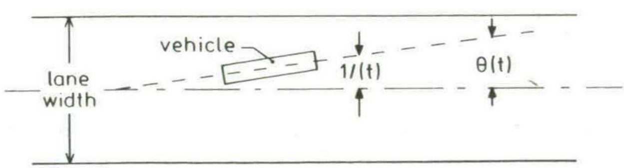


Figure 2. The driving situation.

The other component, $U_n(t)$ also starts from zero but increases without limit as T_d is increased. To determine this component, let us analyze the situation in which uncertainty arises only from lack of knowledge about the lateral position of the vehicle. Assume that during the viewing period the driver is able to center the vehicle within the lane and orient it parallel to the (imaginary) centerline. The driver then has to worry about wandering into the next lane during the period of occlusion of vision. A reasonable rule for the driver to adopt is that σ_y / T_d (the expected value of the rms displacement of the vehicle from the centerline at time T_d) must be less than some level Y_s or that the uncertainty of the lateral position of the vehicle must be less than U_c where the uncertainty is proportional to σ_y / T_d.

We shall now determine σ/T in terms of velocity and occlusion time. The driving situation is diagrammed in Figure 2. Let $\theta(t)$ be the angular orientation in radians of the vehicle with respect to the centerline of the lane. Let $y(t)$ be the lateral distance of the center of the front bumper from the centerline. The lateral component of the vehicle velocity is $dy(t)/dt = V \sin\theta(t)$. Since $\theta(t)$ is typically a small angle,

$$\frac{dy(t)}{dt} = V\,\theta(t) \tag{9}$$

Thus,

$$\frac{d^2}{dt^2}\,y(t) = V\,\frac{d}{dt}\,\theta(t) = V\,\frac{\partial\theta}{\partial x}(t)\,\frac{dx}{dt} = V^2\,\frac{\partial\theta}{\partial x}(t) \tag{10}$$

where $\frac{\partial\theta}{\partial x} = (t)$ is the rate of change of orientation with respect to distance along the road in radians/mile. The value of the lateral displacement at a particulair time T is obtained by double integration of the right-hand expression of Eq. 10 :

$$y(T) = V^2 \int_0^T \int_0^T \frac{\partial\theta}{\partial x}\,(t)\;dt\quad dt \tag{11}$$

In order to compute σ_{y_t}, $E(y^2T)^{\frac{1}{2}}$, we must know something about $\frac{\partial\theta}{\partial x}(t)$ For mathematical ease, let us assume that this function can be described by a wide-band Gaussian random process that has a rectangular spectrum from 0 to ω_1 radian/sec. Let the spectral density be a constant S_θ radian-sec/mile .

The variable $y(t)$ may be considered to be a continuous random process which is obtained by filtering the continuous random process $\frac{\theta}{x}(t)$ with a system that scales by V^2 and performs a running double integral from $T-t$ to T

Hence the relation between $y(t)$ and $\frac{\partial\theta}{\partial x}(t)$, expressed in the frequency domain, is given by the following system function :

$$H(J\omega) = V^2\,\frac{(1-e^{-j\omega T})^2}{-\omega^2} \tag{12}$$

The power density spectrum of $y(t)$, which shall be denoted as S_y can be obtained by multiplying the powerdensity spectrum of $\frac{\theta}{x}(t)$ by the square of the magnitude of the system function. Thus,

$$S_y = S_\theta\,|H(\omega)|^2 = 4\,\frac{V^4(1-e^{-j\omega T})^2(1-e^{+j\omega t})^2}{\omega^4}\,S_\theta \tag{13}$$

$$= 4V^4\,\frac{(1-\cos\,\omega T)^2}{\omega^4}\,S_\theta$$

The average power of the process $y(t)$ is the integral of S_y over the entire frequency range. Thus,

$$y^2(t) = \frac{4V^4 S_\theta}{2_\pi} \int_0^{\omega_i} \frac{(1-\cos\omega T)^2}{\omega^4} d\omega \tag{14}$$

Let a new variable z be defined such that $\omega = z/T$, and assume that negligible error is introduced if the integral is carried to infinity. Then,

$$\overline{y^2(t)} = \frac{4V^4 S_\theta}{2\pi} \int_0^\infty \frac{(1-\cos z)^2}{z^4/T^4} d(z/T) \tag{15a}$$

$$= 4V^4 T^3 S_\theta \int_0^\infty \frac{(1-\cos z)^2}{z^4} dz$$

Since the integral in the above expression is a constant, Eq. 15a can be reduced to

$$\overline{y^2(t)} = 4KS_\theta V^4 T^3 \tag{15b}$$

Assuming that the random process $y(t)$ fulfills the condition of ergodocity (3), we may equate the time average $y^2(t)$ to the ensemble average $E(y_T^2)$ which is identical to $(\sigma_{y_T})^2$. Thus the expected value of the lateral displacement of the vehicle is

$$\sigma_{y_T} \alpha V^2 \cdot t^{3/2} \tag{16}$$

Since the driver's uncertainty concerning the lateral position of the verhicle is assumed to be proportional to σ_y,

$$U_n(t) = K_n V^2 (t)^{3/2} \tag{17}$$

where the constant K_n includes the power density spectrum S_θ and other scaling factors. Substitution of Eqs. 17 and 8 into Eq. 7 shows that the total driver uncertainty is

$$U(t) = H \cdot D(1-e^{(V/D+1/F)t}) + K_n V^2 (T_d)^{3/2} \tag{18}$$

Thus, the driver's rule of behaviour, obtained from Eq. 6, is

$$U(T_d) = H \cdot D\ (1-e^{-(V/D+1/F)T_d}) + K_n V^2 (T_d)^{3/2} \leq U_c \tag{19}$$

* The replacement of t by T as the time index is justified since both $y(t)$ and y_T have been defined to represent deviations of the vehicle T seconds of occlusion

EXPERIMENTAL PROGRAM

An experimental program was devised with two goals. One of these was to provide data with which to test the adequacy of the theoretcial notions previously expressed and to evaluate their utility. The other was to estimate the attentional demand imposed on a driver by various combinations of road, vehicle, and speed, and to explore a wide range of these variables in order to obtain data relating, for example, the radius of a curvature of a highway to attentional demand as a function of speed. Thus, the experiments stand on their own as empirical investigations into the effects of interrupted vision upon driving behaviour.

In order to accomplish the goals of the program, we wished to investigate a broad spectrum of road difficulty or road attentional demand. Accordingly, two kinds of roadway were used. One of these, the "easy one". was I-495 in Massachusetts. The road is essentially straight i.e. the radii of curvature are sufficiently large so that the viewing distance ahead is always large and the lanes are sufficiently wide so that no great precision of steering is required to stay in lane. For the difficult road we chose a closed-circuit, sports-car, racing course, the Bryar Motorsport Park at Loudon, New Hampshire.

Two kinds of experiments were done. One of these involved the use of a constant period of occlusion and a constant observation time, with the driver controlling speed to his maximum. The other used a constant speed and permitted the driver "to look when he wished to." The experimental apparatus was designed to permit both kinds of operation. Driver vision was controlled by a translucent screen which could be lowered over the driver's eyes and through which no road or vehicle detail could be seen. This screen was the pivoting face shield of a protective helmet and was remotely actuated by a pneumatic cylinder and linkage mounted on the helmet itself .

The system provided for a variety of methods of control and safety override. The experimental vehicle was a 1965 Dodge Polara with a number of modifications. Details of roadways, recording and control apparatus, subject population and experimental procedures are given elsewhere *(4).*

There were four experiments using two kinds of road and two procedures :

Experiment 1-I-495, fixed occlusion and viewing time :
Experiment 2-I-495, fixed velocity and viewing time;
Experiment 3-BMP, fixed velocity and viewing time; and
Experiment 4-BMP, fixed occlusion and viewing time.

Table 1 shows the conditions which were experimentally investigated. The entries are for numbers of series each of which consists of a number of runs. A total of more than 550 runs were accumulated.

Experiment 1 - I - 495 : Fixed Occlusion and Viewing Times.

This experiment dealt with the problem of constant viewing time T_2 and constant occlusion time T_d Five subjects were used : C.W.D., J.W.S., D.H.K., W.V.D., and D.C.M. Subject C.W.D. replicated the experiment completely. As a result, six sets of data are availble.

TABLE 1

NUMBERS OF SERIES COMPLETED AT THE VARIOUS CONDITIONS, OVERALL PROGRAM

Condition		Subject D.H.K.	D.C.M.	C.W.D.	W.V.D.	J.W.S.	Total
Experiment 1							
Td	1.0		3			3	6
	1.5	1	3	6	1	3	14
	2.0	1	3	6	1	3	14
	2.5			6	1	3	10
	3.0	1	3	6	1	3	14
	3.5				1		1
	4.0	1	3	6	1	3	14
	4.5				1		1
	5.0	1			1		2
	5.5		3	3			6
	6.0	1		3	1	3	8
	7.0		3	3	1		7
	7.5	1		3	1	3	8
	9.0	1	3	6	1	3	14
Total		8	24	48	12	27	119
Experiment 2							
V	22	2		1	1		4
	25	1		1	1		3
	30	1		1	1		3
	40	1		1	1		3
	50	1		1	1		3
	60	1			1		2
Total		7	0	5	6	0	18
Experiment 3							
V	22	2	1	2	1		6
	25	2		2	1		5
	30	3	1	3	2		9
Total		7	2	7	4	0	20
Experiment 4							
Td	0.5	2		5	2		9
	1.0	2		5	2		9
	1.5	2		3	2		7
	2.0	2		1	2		5
	3.0	2		1	2		5
Total		10	0	15	10	0	35

Based on the preliminary experiments, three values of T_1 were chosen : 0.25 sec. 0.50 sec, and 1.0 sec. The T_1 of 0.25 sec was the shortest practical time which the driver could use. It allowed for at least a change of accommodation, even if not for more than one fixation of field of view ahead. The 0.50-sec. viewing time was apparently long enough to provide nearly all the information needed to drive at any speed, and only a slight increase in velocity was expected to occur with a 1.0-sec. viewing time.

TABLE 2

SPEED IN MPH AS A FUNCTION OF VIEWING TIME[a] AND OCCLUSION TIME FOR SUBJECT C.W.D. (1)

Occlusion Time (sec)	Trial No.	Run No. 1	2	3	4	5	6	7	8	Obt. Mean (n)	Stand. Dev.	Calc.
1.5	2	62	59	62	60	59				60(5)	2	86
2.0	27				no limit							
2.5	9	27	44	45	51	46	50			47(5)	3	54
3.0	19	52	55	57	61	61	60			60(4)	1	45
4.0	28	35	50	45	45	48	40			46(5)	3	34
6.0	21	15	17	17	22	18	16			17(6)	2	21
7.5	22	6	9	13	13	12	13			13(4)	0.5	15
9.0	6	11	6	4	6	4	6			5(5)	1	11
										N= 42		

[a] Viewing time =0.50 sec.

TABLE 3

SPEED IN MPH AS A FUNCTION OF VIEWING TIME[a] AND OCCLUSION TIME FOR SUBJECT J.W.S.

Occlusion Time (sec)	Trial No.	Run No. 1	2	3	4	5	6	Obt. Mean (n)	Stand. Dev.	Calc.
1.0	23	48	53	50	50	49	49	50(6)	2	62
1.5	17	46	48	47	45	45	40	45(6)	2	44
2.0	16	39	39	38	41	38	35	38(6)	2	34
2.5	3	25	32	33	35	32		33(4)	1	28
3.0	12	20	22	20	20	24		21(5)	2	24
4.0	22	19	20	20	19	20	19	19(6)	0.5	18
6.0	6	13	11	12	12	19		13(5)	3	11
7.5	18	11	7	5	5	7	7	6(5)	1	8
9.0	21	4	3	6	5	5	5	5(6)	1	6
								N=51		

[a] Viewing time = 0.25 sec.

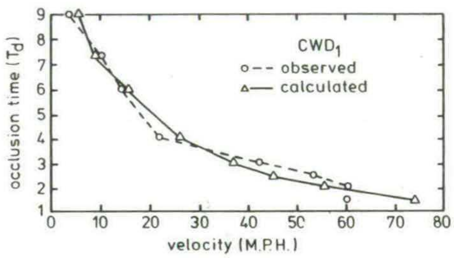


Figure 3. Relationship between occlusion time and terminal speed.

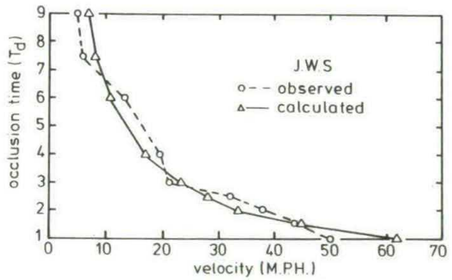


Figure 4. Relationship between occlusion time and terminal speed.

The number of runs made in each series was determined by the stability of the results. The earlier trials tended to be more variable than the later ones as the driver experimented with modes of perceiving, remembering, and controlling. Since each trial, by definition, was terminated only at a limit velocity, and since the limit velocity, by direction, was the "maximum possible". the task of driving was an arduous one. Subjects were relieved after fifteen minutes of driving and, in turn, acted as recorder or experimenter. Trials would take various amounts of time depending on the experience which had been accumulated and the speed involved. The criterion for performance was "adequate driving". The driver was required to stay in lane as he would if he were able to view the road continuously. On no occasion was it necessary for the safety driver to take over control.

The data for C.W.D. and J.W.S. are presented in Tables 2 and 3 and show, on successive runs, the speed reached as a limit by the subject for various T_d The number preceding the tabular entries of mph was the position of that particular combination in the sequence of that subject for that experiment. Thus, the second run made with subject C.W.D. (1) was with a viewing time of 0.50 sec and an occlusion time of 1.5 sec. Runs were continued until, in the opinion of the experimenter, stable performance had been reached. The mean was calculated on the basis of a number of points, never less than 4, which appeared to represent stable performance. The parenthetical entry after the mean is the number on which the mean is based. The standard deviation is computed on the basis of that n and is in mph. The last entry is the speed in mph calculated on the basis of the best fitting solution of the theoretical model.

Figures 3 and 4 show the obtained and the theoretical relationship between T_d and terminal speed. The calculated values are those based on a best mean square fit to the theoretical model. The fit is that which minimizes the variability of the permissible uncertainty, U_c , and allows the other parameters of the equation to vary in order to minimize the variation in U_c . Thus, for each subject, there is a table of the obtained data and of the theoretical points fitted to the curve by a minimization of the squared deviations of the permissible uncertainty, and a graph. Two such sets, for C.W.D. (1) and J.W.S. are shown. The complete data are given elsewhere *(4)*.

The results of Experiment 1 show, as would be expected, that as occlusion time is decreased the maximum velocity which can be achieved by the subject-drivers is increased. With few exceptions, the function is a monotonic relationship and where reversals have occurred they can almost always be identified as being the result of learning.

It is possible to adjust the parameters of the model to provide a good fit to the obtained data in most cases. Table 4 shows the values of the parameters for each subject, for each condition. D, F, and U_c are individual parameters, which, of course, will vary with the subject and with the conditions under which data are taken. The parameters H and K are situational and proportional parameters ; K , in particular, is the same for all, having been set to adjust the general position of the model to correspond to the real numbers obtained in the experimental situation. H presents somewhat more of a problem. As we can see H varies from as high as 34 for subject C.W.D. (2) to as low as 4 for subject J.W.S. Whether this represents a different attitude toward the roadway is not clear. Subject J.W.S. uniformly had a very small D which suggests that his performance was based largely on the information relatively close to the vehicle. Subject C.W.D. had a somewhat higher D factor and, in general, a much larger H. In this respect, it must be noted that subject J.W.S. was, in general, a more cautious subject that C.W.D. or D.C.M. In particular U_c -- the amount of uncertainty which the driver will permit to be accumulated in bits -- is consistent and small for subject J.W.S., and larger and more variable for the other two subjects.

An impressive feature of the data is the fact that, with the exception of subject J.W.S. an occlusion time 1.0 sec resulted in no limit speed -- at least no limit within the speed limit of the highway. The only exception to this finding occurred with a, T_1 of 0.25 sec for subject C.W.D. (1), and on one trial with a T_1 of 0.5 sec. At the other extreme, subject C.W.D. (2) drove with complete control with no tendency to deviate from his path or to exceed the limits of his lane at speeds in excess of 70 mph with 1.0-sec looks at the road separated by intervals of 4.0 sec of complete occlusion of vision.

TABLE 4

MODEL PARAMETERS FOR VARIOUS T_1

Subject	Parameter	T_1 0.25	0.50	1.0
C.W.D.(1)	D	0.26	0.42	0.32
	H	14.0	12.0	20.0
	F	6.0	10.0	7.5
	K	0.0002	0.0002	0.0002
	U	3.13	3.76	5.22
J.W.S.	D	0.20	0.18	0.20
	H	6.0	6.0	4.0
	F	9.5	6.0	5.0
	K	0.0002	0.0002	0.0002
	U	0.99	1.07	1.13
D.H.K.	D		0.50	
	H		18.0	
	F		3.5	
	K		0.0002	
	U		4.93	
C.W.D. (2)	D	0.22	0.46	0.20
	H	12.0	34.0	26.0
	F	8.0	5.0	7.0
	K	0.0002	0.0002	0.0002
	U	2.57	13.47	8.24
D.C.M.	D	0.30	0.30	0.30
	H	20.0	20.0	12.0
	F	4.5	5.0	3.5
	K	0.0002	0.0002	0.0002
	U	7.16	5.96	4.29
W.V.D.	D		0.32	
	H		24.0	
	F		4.5	
	K		0.0002	
	U		6.99	

Experiment 2-I-495: Fixed Velocity and Viewing Time, Voluntary Control of Occlusion.

This experiment dealt with the situation where speed and T_1 are fixed and T_d is under the voluntary control of the driver. As described earlier, the driver has available a switch (to be operated by the left foot) which initiates an observation time of fixed duration equal to T_1. The purpose of the experiment was to obtain something more analogous to a point-to-point index of the attentional demand placed upon the driver by the highway. A fixed T of 0.5 sec was used. The subject was fitted with the helmet and given preliminary familiarization with the operation of the foot switch and the viewing time.

TABLE 5

MEAN VOLUNTARY OCCLUSION TIME AS A FUNCTION OF SPEED, EXPERIMENT 2

D.H.K.		W.V.D.		C.W.D.	
V	$\overline{T}_d$	V	$\overline{T}_d$	V	$\overline{T}_d$
60	1.48	60	1.84		
50	1.66	50	2.50	50	2.21
40	1.75	40	2.82	40	2.42
30	2.10	30	3.19	30	3.25
25	2.26	25	3.95	25	3.57
22	3.63	22	3.98	22	3.64
22	2.60				

The subject steered into the right-hand lane of the road and accelerated with the visor up until he reached the preset speed. When the driver had provided an adequate sample of behaviour, i.e., some 60 to 120 operations with the visor in the course of 5 minutes, the experimental run was terminated with the visor's being raised, a new speed selected, and the experiment repeated.

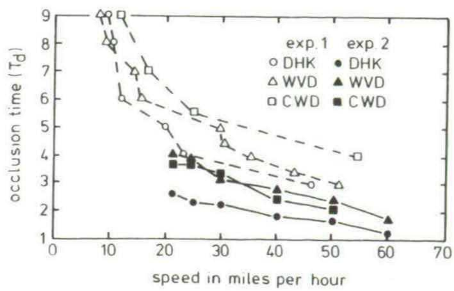


Figure 5. Viewing time 0.5 sec.

The average T_d 's chosen by the various subjects for the various speeds resulting from the process analysis are given in Table 5. It can be seen that there are no reversals of the functional relationship between V and the mean value of T_d chosen by the driver. The longer times obtained for subject D.H.K. were the result of his not adhering to the rule of error-free performance on the first 22-mph run. When the instruction was reiterated the shorter times were obtained. The results are shown in Figure 5 with the functional relationships between V and T_d for the same subjects obtained in Experiment 1 presented for comparison.

For all three subjects, it can be seen that the voluntary control T_d 's are lower than would have been expected if this technique resulted in a simple replication of the results of Experiment 1. From the subjective point of view, the technique of Experiment 2 provides the driver with a more immediate control over his perceptual environment. The response time of the visor to the depression of the foot switch is virtually instantaneous. On the other hand, the vehicle, when traveling at a fairly high speed, is relatively slow to respond to minor changes of accelerator or brake pressure which might be applied in Experiment 1. The assumption of a constant limit speed in Experiment 1 is predicated on the idea of a constant H for the highway and this is almost surely not the case. The variability which is evident in the voluntarily controlled T_d 's shows that the driver perceives the highway as possessing a variable H . Consequently, there is variation in the rate of uncertainty generation during occlusions. This suggests that the technique of Experiment 2 will be a more useful one for evaluating the attentional demand placed upon the driver by traffic situations or by different vehicles.

Experiment 3-Bryar Motorsport Park : Fixed Speed and Viewing Time, Voluntary Control of Occlusion.

This experiment dealt with the problem of constant viewing time, T_1 , and constant velocity, with subject control T_d. It was thus identical with Experiment 2 but done on a very different road. Bryar Motorsport Park is 1.6 miles of well-paved and banked roadway with ten turns which vary in radius from vitually straight to "hairpin".

As was seen in Table 1, a total of 20 trials was made, with each trial consisting of a number of laps. Each lap produced a record which was analyzed separately. Our interest was in the functional relationship between the voluntary period of occlusion and the radius of curvature of the track. Since no other vehicles were permitted on the track and there was a sufficiently long period of familiarization, the only residual uncertainty would be that associated with steering around curves and maintaining the vehicle properly within the lane. Accordingly experiments were done at three constant speeds : 22, 25 and 30 mph. A speed of 35 mph is not beyond the limits of the vehicle but would present serious control problems in the event of error. For this reason, no higher speed attempts were made. (These were reversed for Experiment 4, which dealt with constant T_1 and T_d and allowed the driver to vary the speed at will). A sample record (Figs. 6 and 7) shows the interval between observations as a function of position along the track. The data on a point-to-point basis are jagged, due to the discrete nature of the performance. Accordingly, running averages of 3 were made and are plotted on the same graph. The numbers of the record indentify the five significant curves of the track.

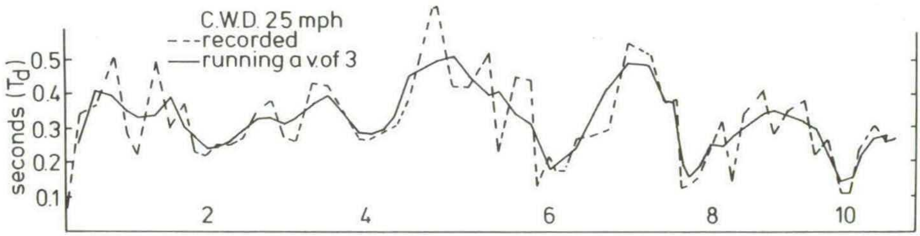


Figure 6. Interval between observations as function of position on track.

It can be seen that, in general, the higher the speed, the shorter the interval between observations. For a sufficiently low speed, we might assume that the driver will behave as he does on the superhighway. That is to say, the interval between observations would be nearly independent of the radius of curvature on the roadway. At sufficiently high speeds, on the other hand, the driver might be unable to maintain adequate steering, given a roadway that curves as it does, even with continuous viewing. At intermediate speeds, on some parts of the track the roadway will be effectively a straight road (even though in fact it may be curved), and on other parts of the track the radius of curvature might be so small as to demand more frequent viewing. We should not expect, therefore, to find a simple relationship between the speed at which the track as a whole is traversed and the total number of observations which might be made of it at any speed, since in a sense each increment of speed merely increases the part on which more frequent observations must be made rather than requiring that this be done over the whole length of the track.

If we examine the performance of W.V.D., for instance, we find that the total number of looks increases slowly with increasing speed. But there is surprisingly little difference as the speed changes from 22 to 30 mph. Thus, it would appear that the major factor which induces a new look at the road is that of distance traversed rather than either time or speed per se. Thus, at 22 mph 74 looks were taken at the track; at 25 mph 79 looks; and at 30 mph 80 looks.

We find that our estimate of the distance traveled between observations is 83.6 ft for 30 mph, 87.7 ft for 25 mph, and 100 ft for 22 mph. Since, presumably, the information content of the roadway is invariant and the U_C of the driver is likewise invariant, the variation in distance traveled must correspond to an increase in the noise power generated by the vehicle itself. In other words, at the higher speeds, the precision of the driver's steering and aiming, as well as the residual uncertainties in the steering mechanism, become more important and require more frequent observations.

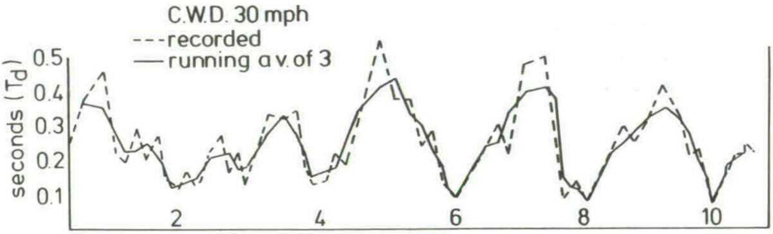


Figure 7. Interval between observations as function of position on track.

For subject C.W.D., the mean number of looks taken at the track is 69 for a speed of 30 mph, 68 for 25 mph and 67 for 22 mph. The corresponding distances traveled between observations for these same speeds are 97, 102 and 102 ft for 30, 25, 22 mph respectively. Thus, although there is less variation for the three speeds for this driver, the same general trend can be seen.

A factor which has not so far been considered in our discussion of Experiment 3 is the "size" of the turns or total amount of direction change. Thus, turn 2 and turn 6 involve very nearly 180 deg of direction change, turn 8 somewhat less, turn 10 still less and turn 4 less again. Turn 3 has a fairly large radius of curvature that involves a directional change of about 90 deg. Turns 5,7 and 9 have the largest radii of curvature and involve changes of direction of only 15 deg. , approximately. The more directional change there is, the more interference with vision ahead and the more demanding of attention from the driver. Thus, turn 3 has a radius of curvature approximately the same as curve 4, but involves only about half as much directional change, and it can be seen that curve 4 elicits observations from the driver in general much more often than does curve 3. Presumably, this elicitation is a result of the limitation of forward view rather than the radius of curvature per se.

Curve 10 is a special case due to the departure of the exit roadway immediately prior to its entrance. In addition, there are bridge railings which constrain, to some extent, the freedom of the driver to approach the edge of the road on the outside coming into curve 10.

The remaining data for all subjects, although not presented here, have been analyzed sufficiently to show that they conform to the same general pattern.

An effort was made, on the basis of maps of the track, to determine the radius of curvature of the various curves. These were then ranked according to radius, the smallest number being the smallest radius. Similarly, the interobservation interval utilized by the subject was also ranked, the smallest number being given to the smallest interval. In this way, we were able to get an estimate of the extent to which there is agreement between the attentional demand as measured by the interval between observations and the radius of curvature of the road.

Although there is general agreement, there are factors which prevent us from arriving at a firm conclusion. First, in addition to radius of curvature and total extent of directional change which have been noted, there is also the effect of road width which controls the actual path which can be taken by a driver in negotiating a curve. A driver driving on a track attempts on each curve to negotiate the curve with a path of maximum constant radius of curvature. This path of "maximum constant radius" (MCR) minimizes the degree of control activity required and also permits the fixed speed to be most easily maintained through the curve. The line of MCR is a function both of the radius of curvature of the track itself, the width of the track, and the amount of direction change which the curve involves. Quite clearly, a circular track would have an MCR precisely equal to the radius of the outer boundary of the track. A segment of that curved road resulting in a change of perhaps 10 deg and connected by straight segments before and after will have an MCR which would be many times larger and which would be a direct function of the width of the track.

Curve 10 appears to be the most demanding even though its radius of curvature shares only the second rank with curve 6. However, the constraints on the entrance to curve 10 mean and the MCR which the driver can attain is small with the consequence that the "functional radius" of this curve may in fact be smaller than the "functional radius" of any other curve. Thus, there is not a monotonic relationship between radius of curvature and attentional demand, such that the attentional demand is always less when the radius of curvature is larger, but rather the constraints on paths through the turn must also be taken into account. Of course, in dealing with highway curvature where cars are by custom or by law required to stay in lane, the curve taken by the car has a radius more nearly equal to that of the road and thus a more precise relationship could be obtained. On exit roadways from superhighways as well as on roads where visibility around curves is good due to the flatness of the terrain, there is an almost unavoidable tendency on the part of the drivers to cut curves and in this way enlarge the radius of curvature with which they negotiate the curve. This is observed, also, in open highway driving where there is little opposing traffic.

However, and this finding is more general, by comparison of the voluntary intervals between viewing on the superhighway and those obtained on the race track, it can be seen that when the average H of the road increases, the interval between observations at a given speed tends to decrease. We have, of course, only two points and we do not have any immediate estimate of where these points lie, since the fitted curves for the various drivers are best fitted by different H's. Whether H is therefore a demonstrable external physical variable or one which is a compound of psychological and physical variables is yet to be determined. The extent to which a driver is familiar with the very minor aspects of the road as well as with its statistical structure (that is, the probability of intrusion, etc.) probably modifies for him whatever basic H or uncertainty exists in the road purely because of its physical nature. Thus, the timorous, unfamiliar driver sees the same road as possessing a larger H than the man who has traversed it many times. Presumably, this would be one source of the variation in the value of the fitted parameter H of the various subjects in Experiment 1. More detailed examination of the visual stimuli present at various points along the track as well as those which exist along the highway is needed.

Experiment 4-Bryar Motorsport Park : Fixed Viewing and Occlusion Times, Voluntary Control of Vehicle Speed.

The fourth experiment dealt with the problem of constant viewing time T_1 and with the constant occlusion time T_d while the driver retains control of the vehicle speed. The experiment is thus identical with Experiment 1 but done at Bryar Motorsport Park. The operating procedure was identical with that described in Experiment 3. After familiarization runs the experimenter in the left rear seat initiated a run by setting a particular T_d with a fixed T_1 of 0.5 sec., and started the apparatus. The driver then circled the track driving as rapidly as possible and varying his speed from point-to-point along the track. Errors were not permitted; that is to say, the driver could not leave the road or cross the white line at the edge for any reason. If he did so, the run was terminated and another run begun. At the conclusion of three satisfactory runs the visor was opened by the experimenter at the finish line and the driver would rest briefly before initiating runs with a new T_d

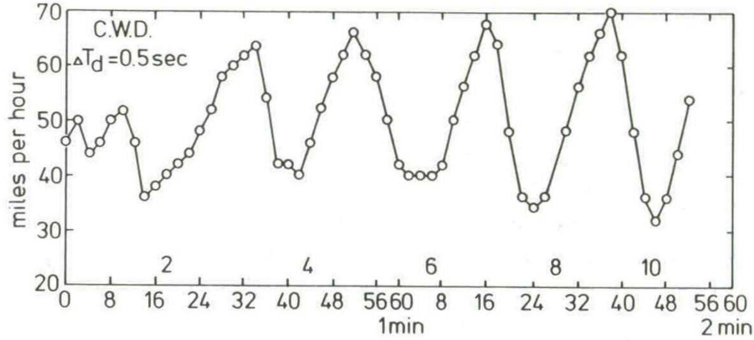


Figure 8. Speed attained as function of position on track

Thirty-five trials were made at T_d 's of 0.5, 1.0, 1.5, 2.0 and 3.0 sec., by subjects D.H.K., C.W.D., and W.V.D. Each trial consisted of a number of runs as previously described. Each lap produced a record which was analyzed separately. Our interest, of course, was in finding a functional relationship between the speed adopted by the driver and the radius of curvature of the track for each given T_d Since no other vehicles were permitted on the track and there had been a sufficiently long period of familiarization, the only residual uncertainty would be that associated with steering on the curves and maintaining the vehicle within the lane. In addition, the available view ahead was a function of the radius and of the total directional change of each curve as indicated in the discussion of Experiment 3.

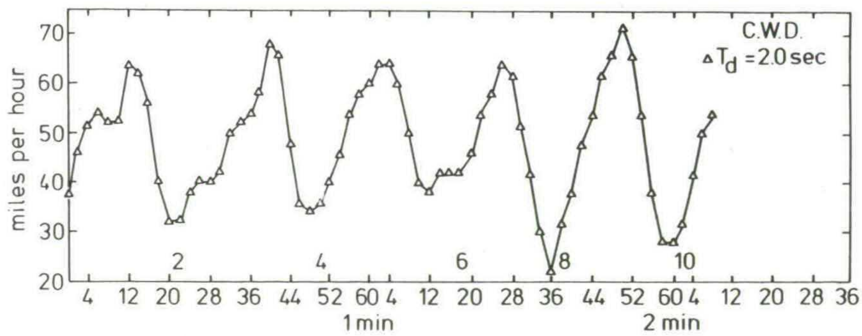


Figure 10. Speed attained as function of position on track.

The data were obtained in the form of strip chart recordings of speed. The general form of the functions is similar to that for the constant velocity an voluntary control of visor time. In other words, if the driver is faced with a greater attentional demand at a constant speed he must look more often and similarly, if his information intake is limited, he varies his speed accordingly. The data for one subject, C.W.D., for T_d of 0.5 and 2.0, are shown in Figures 8 and 9. There is a consistent reduction of speed on the curves with increasing T_d It must also be noted that the increased H of the track produced speed-limited runs even with occlusion times as short as 0.5 sec. It would have been possible for the vehicle to accelerate to higher speeds than those which were in fact reached, as was evidenced by the performance of the same driver with the same vehicle with no occlusion, striving for maximum circuit speed of the track.

The increased H of the track as compared with the Interstate highway resulted both from the curvature and from the limitation of the viewing distance ahead, and markedly reduced the speeds which the driver could attain with occlusion times in the region from 1 to 3 sec. It is difficult to identify on the speed records the odd numbered curves, since the reduction in speed for these was very small, if present al all. However, as is evident from Figures 8 and 9, the salient features of the record are those corresponding to the even numbered curves and it is in these turns that a reduction in speed with increasing T_d is most evident. The records for the other subjects show the same general trend. In order fully to understand the relationship which are implicit in these data it will be necessary to obtain the same information about maximum constant radius as is required to deal with the data for Experiment 3. As a result, a full analysis and interpretation of the results of this experiment must await the availability of these additional data.

GENERAL DISCUSSION AND RECOMMENDATIONS

The general purpose of the experiments was to determine empirically certain relationships between characteristics of the road upon which a car is driven, the amount of time a driver has to look at the road, the interval between such observations, and the speed at which he drives. Experiments 1, 2, 3 and 4 attempted to do so for two different classes of roads and for two different modes of operation of the experimental apparatus. The results indicate, as would be expected, that the less frequent the observations, or the shorter the period of observation, the slower will be the speed that the driver can maintain, and, conversely, that the greater the level at which the speed is fixed the more often the driver must look at the road. In addition, the difference between the roads appears as a modifier in that the more complicated road results in a lower speed at any constant viewing and occlusion times, and results in shorter occlusion times for any constant speed and viewing time. The data with which to express the functional relationships among all of these variables have been obtained and subjected to partial analysis.

Another part of the program was aimed at testing the adequacy of a theoretical model which described the behaviour of the driver in terms of information processing and uncertainty accumulation. The data for Experiments 1 and 2 provide an opportunity to verify the adequacy of the model. Figures 3 and 4 show the extent to which the model fits the observed data. Table 4 gives the parameters of the model for the various subjects for the data for Experiment 1. Because of time and cost limitations, no effort was made at this time to apply the model to the data of Experiment 2. Instead, a simple comparison of the latter data has been made with those of Experiment 1 to show that with voluntary control of occlusion by the subject the times which he generates himself are somewhat shorter for any given speed than those which permitted that same speed in Experiment 1. Experiments 2 and 3, both of which permitted voluntary control of occlusion time by the driver, suggest by their results that this technique is a useful one for measuring the attentional demand of a driving situation. Figures 6 and 7 show quite clearly that for various constant speeds, the driver must look more frequently as he enters and passes through the various curves on the track, and his increased frequency of observation is not independent of the nature of the curve itself. Thus, we believe that this technique will allow an objective measure, based on driver behaviour, of the attentional demand of any segment of a road.

Data about the view ahead must be obtained by field measurements both on I-495 and at Bryar Motorsport Park and the curve of maximum constant radius must be calculated for the various turns. When this has been done, it should be possible to establish the parameter H for the various parts of the track, and to calculate more precisely the driver uncertainties, distance constants, and forgetting time constants which will be needed to provide a crucial test of the ability of the model to predict the behaviour of a driver on some new road.

If the model can make such predictions, then the identification of these parameters D, F, and U_C might permit a preliminary classification of drivers in terms of skill level. This, in turn, suggests a means of identifying those drivers who may be potentially "accident prone". The use of trained drivers as measuring instruments, using the voluntary occlusion technique, may facilitate the identification and quantification of hazardous or excessively demanding road configurations or vehicle characteristics.

ACKNOWLEDGMENT

The work reported was initiated under contract from the U.S. Bureau of Public Roads and continued by the principal author under a grant from the Accident Prevention Division of the U.S. Public Health Service.

REFERENCES

1. Gordon, Donald A. Experimental Isolation of Driver's Visual Input. Public Roads, Vol. 33, No. 12, pp. 266-273.

2. Sheridan, T.B., Merel. M.H., Kreifeldt, J.G. , and Ferrell, W.R. Some Predictive Characteristics of the Human Controller. Progress in Astronautics and Aeronautics, Vol. 13, Academic Press, Inc., New York, 1964.

3. Davenport, W.B. , Jr. , and Root, W.L. An Introduction to the Theory of Random signals and Noise. McGraw-Hill, New York, 1958.

4. Senders, J.W. , Kristofferson, A.B. , Levison, W.H. , Dietrich, C.W. , and Ward, J.L. An Investigation of Automobile Driver Information Processing. Bolt Beranek and Newman Inc. , Final Report No. 1335 under Contract No. CPR-11-0958 for U.S. Bureau of Public Roads, April 1966.

APPENDIX C

Axiomatic Models of Workload

A. A SIMPLE STEADY STATE MODEL OF WORKLOAD

1. Workload is a meaningful concept only in the context of a well-defined task which must be performed to a stipulated criterion.

   a. - def. a 'well-defined task is one for which specific actions on the part of a human operator can be identified and quantitative criteria for each action stated.

   b. - def. workload is a non-accessible, hypothetical, unidimensional, internal variable over the range of 0.0 - 1.0 in any human operator.

2. Workload increases as the demand of the task increases.

   c. - def. Demand of a task is the list of actions and their performance criteria.

3. Workload decreases as the capacity of the human operator increases with respect to that task.

   d. - def. Capacity of a human operator is the limiting level of performance on any action which is part of a task.

4. By 2. and 3. workload may be considered to be the ratio of Demand to Capacity with respect to action. Thus $L_i = \frac{D_i}{C_i}$ or the workload of action is the ratio of the demand of action i to the capacity to perform action i.

5. Actions requires by a task may be classified into elemental types.

   e. - def. An elemental type is a non-reducible action.

6. All actions required by a task are either elemental types or are reducible into elemental types.

7. The non-reducibility of elemental types implies independence of elemental types.

8. By 7. , the demand of a task can be mapped into an N-dimensional space whose orthogonal axes are the elemental types of the task and are also the capacities associated with each elemental type by def. d.

9. The N-dimensional space of elemental type demands and capacities may be transformed into an N-dimensional space of elemental type workloads. The axes of the new space are not limited to the range 0-1.0 ; the mapping consists of dividing elemental type demands by elemental type capacities.

10. The total workload of the task may be considered to be the vector sum of the elemental type workloads.

11. If a task demand consist exclusively of a repetitive elemental type, the load placed on the human operator will by $N.L._i$ where N is the number of actions of type i . In all other cases the total workload $L = \left[\sum_{i=1}^{N} (M_i L_i)^2 \right]^{\frac{1}{2}}$ which is the length of the vector composition of the elemental type work loads. N is the number of elemental types and M_i the number of each elemental type i .

12. If any elemental type workload exceeds 1.0, the human operator cannot meet the demand of the task. If the vector task workload exceeds 1.0, the human operator cannot meet the demand of the task.

13. Summary of simple steady state model :

    Workload of a well defined task is the length of a vector in a space defined by *N* elemental type workload axes. On each axis is plotted the workload corresponding to that elemental type. Each such elemental workload is computed by taking the ratio of elemental type demand to elemental type capacity.

B. SIMPLE NON-STEADY STATE MODEL OF WORKLOAD

1. In any real task the numbers and type of elemental type actions which compose the demand of the task vary as a function of time.

2. As a consequent the vector sum of the elemental workloads varies in time in both length and angle relative to each elemental type axis.

3. In any real human operator the capacity to perform any elementary action varies as a function of time.

4. As a consequence, the vector sum of the elemental type workloads varies in time in both length and angle relative to each elemental type axis.

C. THE LOAD COMPONENT GENERATED BY INTERNAL SWITCHING

1. Subjective report strongly indicates the existence of a workload component generated internally to the human operator.

2. The internally generated workload component may be considered to be orthogonal to and similar to the elemental type workloads demanded by the task.

3. The internally generated workload increases monotonically with increasing number of elemental type workloads demanded by a task.

   Hypothesis : The workload generated internally is $L_N + 1 = K.N$, where N is the number of elemental type workloads and K is the workload associated with the switching from one elemental type to another.

4. It follows that the workload vector is in a space of $N+1$ dimensions and will increase with increase in N.

D. THE LOAD COMPONENT GENERATED BY EXTERNAL SWITCHING

1. Subjective report and direct observation require that a workload component is generated by switching of attention from place to place and from signal to signal.

2. Switching of visual directions of attention requires some minimum time τ_1.

3. Switching from sensory mode to sensory mode requires some minimum time τ_2.

4. From the task description the number of loci of visual attention can be calculated and the number of intersensory switchings can be estimated.

5. The time lost in switching is then $T_L = k_1\tau_1 + k_2\tau_2$ where k_1, k_2 are the numbers of switchings which the human operator must perform.

Hypothesis : Total capacity but not elemental type capacity is reduced by T_L/T when T is the total time available to perform the task.

E. SWITCHING LOADS AND TIME COST

1. In external switching as in an extended monitoring task, the principal demand is for the switching act itself.

2. Therefore Time Demanded is the measure of demand and Time to Perform is the measure of capacity.

3. The workload for an externally limited task is $\frac{T_d}{T_c}$ where T_d is the time demanded and T_c is the time required by the human operator to perform.

4. The relationship of timing of elemental type demands relates to workload.

5. If actions are demanded independently, two or more actions may be required simultaneously. There are two consequences.

   1) Priorities accumulate and force uneconomic switching.

   2) Human operators will fail to perform required actions until after an undersirable delay.

6. A totally self-paced task can never have a workload greater or less than 1.0. No overload can exist in either transient or steady-state conditions.

7. Because of the load cost of internally generated switching load, the adaptive human operator will minimize the switching load by aggregating demanded actions into groups of the same elemental type if priorities permit.

8. Two identical streams of actions, one self-paced and the other externally-paced will generate different load levels and different physiological response levels. A behaves in a self-paced mode. B receives as demanded action whatever A has done and is therefore externally paced.

9. Some elemental capacities are single channel in nature. The elemental type load is therefore always 0 or 1.0 with the average load reflecting the relative proportion of time the load is 1.0.

10. For such elemental types queueing models are required for the calculation of load as a function of time and for the estimation of the probability of transient overload.

F. MOTIVATION AND EFFORT

1. Motivation (Effort) controls the fraction of capacity (total) which a human operator will commit to a task.

2. Differing payoff systems and task designs will alter capacity and, therefore, the load imposed by a task.

3. High motivation allows a high fraction of capacity to be used. This results in improved performance. Low motivation produces low available capacity and permits overload to occur at low levels of demand.

   Hypothesis : Available capacity is product of Motivation and Capacity $C_A = M \cdot C_T$ where M is a 0-1 variable.

APPENDIX D

A Queueing Model of Monitoring and Supervisory Behaviour

1. Introduction

In many modern industrial processes the human beings involved have been removed from direct control and placed in a supervisory or monitoring position. Typically, a complex continuous process is represented to the supervisor by some large number of displays of information relating to the state of the plant. The information may be displayed in a large number of different forms. There may be displays of the point values at each moment in time. There may be displays of recent past history as well as current values. In still others there may be remote historical data which serve as a base against which to judge the significance of current and immediately past values of the variables.

The design of workplaces for such supervisory activity must be based on a number of factors. Economic and physical space limitations obviously influence design. In addition there is a necessary logic which relates the characteristics of the displayed variables to the characteristics of the human supervisor. Thus 'Human Factors Engineering' usually concerns itself with the design of individual displays and, in general, there is a sufficient understanding of the design of satisfactory information displays and of the logical organization of large numbers of displays to make clear, even to an only moderately skilled monitor, the displayed data and the functions of the displays.

Two different but related problems which remain are : the calculating of the manning requirements, and the calculating of the reliability of the man-machine interaction. One would prefer to solve these before a plant is built, and surely before it is set into operation.

1.a Manning Requirements

Any valid way of estimating manning requirements would make possible more precise estimates of operating costs. It would facilitate estimates of the effort likely to be involved in recruiting, selecting, training, and evaluating operators, and thus the cost of maintaining a sufficient cadre of trained operators to keep a plant operating continuously.

1.b Man-machine Reliability

In general, the reliability of complex systems is calculable with fairly standard analytical techniques. However, when human beings are introduced into systems, it has usually been more difficult to estimate the probability of failure of those parts of the system in which the human operator plays a significant role. It is intuitively clear that reliability goes down when the load on a human operator goes up, at least when the load is already high. Hence, the relation between the two problems.

2. The Behaviour Model

Most monitoring tasks can be described quite simply as situations in which the human monitor observes one indicator at a time and progressively looks at, or attends to, the various instruments and indicators on which information is displayed.

In addition, of course, there are emergency indicators which are not continuous functions of plant state, but artificially dichotomized functions, such that when certain continuous variables approach critical boundaries, an emergency or alerting signal might be triggered. Our concern here, however, will be only for steady-state, normal operating conditions in which emergency signals do not arise.

One way of conceiving the operator's task is to imagine that each supervisory monitor consist of a 'service channel'. That is to say, various instruments 'come' one at a time, to be served, much as customers approach the teller in the bank or the check-out counter in the supermarket. Thus, it would seem appropriate to consider the queueing characteristics of instruments. Since instruments do not in fact 'arrive'' at the visual system for service in the form of an observation but are serially fixated by the monitor, it is necessary to construct a behavioral model for estimating both the inter-observation intervals and the durations of the observations for any instruments.

Instruments are designed predominantly for foveal viewing. In general, it is difficult, if not impossible, to get information from them if they are not directly fixated. It can rarely be the case, therefore, that an event which happens on a particular instrument elicits the response of the monitor toward that instrument. Instead, internal events in the monitor must be responsible.

2.a Uncertainty Model.

One model for such internal events is that of observer uncertainty about the nature of the data presented on the instrument. In one formulation, when the uncertainty rises above some maximum permissible uncertainty, the probability of fixation of that instrument becomes high. In a situation in which 'overload' did not exist, one would imagine that the instrument about which the operator was most uncertain would generally be fixated next. If the rates of growth of uncertainty of the instruments are widely different, a somewhat more complex algorithm can be postulated involving some weighted summation of uncertainty and rate of change of uncertainty. The latter is brought in because an observer's algorithm for choice of next point of regard may include calculation of the uncertainty of distribution one fixation hence.

The uncertainty that an observer has about a system variable depends upon two things which are related to the time since its last viewing. One of these, which is the same for all signals, is the decay of information stored in short-term memory. The second kind of uncertainty stems from variations in the system being monitored. Both the memory decay and possible plant variation generate monotonically increasing uncertainty about each variable from the moment of past observation. These two kinds of uncertainty, both increasing with time, and presumed to be linearly additive, will eventually exceed a threshold of permissible uncertainty, and elicit a fixation upon the instrument about which uncertainty is maximum or for which some weighted sum of uncertainty and uncertainty rate is maximum.

Each instrument, then, will be observed at some interval drawn from a distribution of intervals unique to that instrument, or, if there are sets of identical instruments, that class of instrument.

2.b Probability Model

The probability model is that which was described in Senders et al (1). The task of the observer is defined as that of detecting values of any of the signals which exceed arbitrary limits. The limits are assumed to be characteristic of the individual variables being monitored. The model calculates the interval between observations as a function of two alternative strategies on the part of the monitor. In one strategy, the observer samples when the probability that the signal exceeds the limit is a maximum; in the other, he samples when the probability itself exceeds some threshold probability. These different goals yield different values of interval between samples. In addition, the interval is itself a variable since it depends on the value of the signal last observed as well as upon the bandwidth and memory of the monitor.

2.c General Comment

The earlier work did not include the effect of forgetting in its solution but that omission is easily overcome. The critical point is that the two models briefly described here : uncertainty-based sampling, and probability-based sampling yield widely differing values of the interobservation interval. Consider the case where the observed value of a signal is very near the limiting value. The probability model predicts a next observation almost immediately after the preceding one since the probability that the signal will exceed the limit is very great. The uncertainty about the value of the signal is still, however, at a very low value, and the uncertainty model would predict a sample at a much later time quite independently of the value previously observed.

3. The Sampling Models.

In all cases the calculations of the intervals between samples is a function of the bandwidth of the signals. If forgetting were assumed absent, the mean intervals would be in strict proportion to the bandwidths, although the actual values would vary with the model. The uncertainty model, if the maximum permissible uncertainty is the same for all signals as assumed, yields fixed intervals. The probability model yields distributions of intervals. Since the queueing model which follows requires only the mean intervals between observations, we will not here consider the distribution functions.

3.a The Uncertainty Model

The externally generated uncertainty (on the observer's part) about the value of a signal is the information generated by that signal during the period of non-observation. For bandlimited Gaussian signals, the information rate, and therefore the uncertainty growth rate is :

$$\dot{U}_s = 2W \log_2 \sigma \sqrt{2\pi e} \text{ bits per second,} \tag{1}$$

where W is the bandwidth in Hz and σ is the average power of the signal. This rises to a maximum of :

$$U_{max} = \log_2 \sigma \sqrt{2\pi e} \text{ bits,} \tag{2}$$

which is the information of a Gaussian distribution after $\frac{1}{2}W$ seconds, the Nyquist interval.

Added to this is uncertainty generated by the forgetting process. If one assumes that the original perception is correct, i.e., within some acceptable limits of error, and that the memory trace becomes contaminated with Gaussian error with the passage of time, then the uncertainty will be :

$$U_f(t) = \log_2 \sigma \sqrt{2\pi e} \quad (1-e^{-.1t}) \text{ bits} \tag{3}$$

assuming a time constant of forgetting of .1 second. The total uncertainty $U_t(t)$ will therefore be :

$$U_t(t) = U_s(t) + U_f(t) = 2Wt \log_2 \sigma \sqrt{2\pi e} + \log_2 \sigma \sqrt{2\pi e} \ (1-e^{-.1t}) \tag{4}$$

The maximum value of the sum will still be the value of equation U_{max}. The observer will sample when the sum reaches some threshold amount U_p. If there were no forgetting of the previously read value of the signal, then the durations for the various signals would be in strict proportion to their bandwidth. What actually happens is that very slowly varying signals will be sampled more often than on the basis of bandwidth alone. See figures 1 and 2.

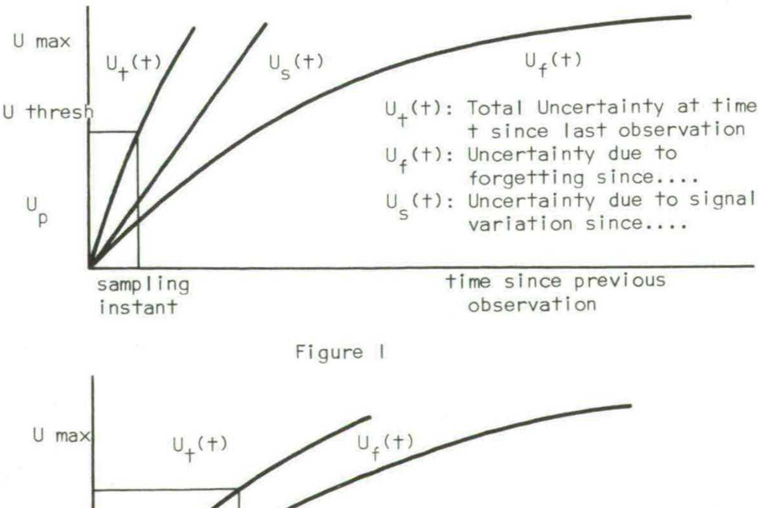


Figure 1

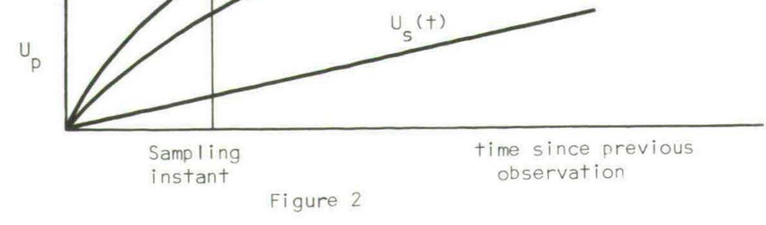


Figure 2

3.b The Probability Model

The analysis shown in reference I yields the following results.

Since the sample interval is a function of bandwidth, the solutions have been expressed in terms of the value of the autocorrelation function. For any bandwidth, these can be converted back into time.

For the case where the observer samples when the probability of exceeding the limit is maximum, the mean value is :

$$\overline{r_m} = \frac{\sigma}{L\sqrt{2\pi}} \left(1 - e^{-\frac{L^2}{2\sigma^2}} \right) + \frac{1}{2} \left[1 - \phi \left(\frac{L}{\sigma} \right) \right] \tag{5}$$

where ϕ is the normal probability integral, L is the arbitrary limit and σ^2 is the variance of the amplitude density distribution of the signal.

For the case where the observer samples when the probability of exceeding the limit itself exceeds some probability threshold, the mean value is :

$$\overline{r_+} = \int_0^L r_T \, r(Y) \, dY + \frac{1}{2} \left[1 - \phi \left(\frac{L}{\sigma} \right) \right] \tag{6}$$

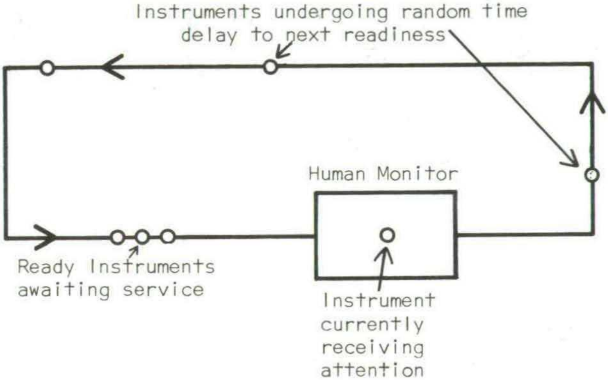
To the results of both of these for any specified function of a given bandwidth limit value, and variance, must be added the same forgetting term used earlier. However, the value of the forgetting variance would be added to the variance resulting from the diminishing value of the autocorrelation function with increasing time since previous observation.

4. The Queueing Model.

Posner and Bernholtz (2) have considered a broad class of closed queueing networks in which customers possessing different characteristics circulate among a number of service stations. A particular subclass of those models studied appears applicable in the present context. Specifically, assume that the system being considered consists of m different instruments which arrive for service (for observation by a single human monitor) on a first-come-first-served basis (see Figure 3). That is, when an instrument requires observation, (because its value may be at the threshold) it may be viewed as moving into a queue to await attention by the observer. The service or attention time requirement for instrument type $(j = 1, 2, \ldots, m$ is assumed to be an exponentially distributed random variable G_j with mean $1/\mu_j$.

After completion of service, instrument j (indicated by 1_j) undergoes a random delay having mean U_j before it is again ready to be observed, and enters the queue. Note that in practice only the mean delay to readiness is required for this model, the results being independent of the shape of the delay distribution.

Figure 3



It is assumed throughout that the process has been ongoing for a considerable time so that steady state conditons prevail, and all limiting distributions to be defined exist.

Let $P(i_1,1_2,\ldots,i_m)$ be the joint stationary probability that i_j instruments of type j $(j=1,2\ldots,m)$ are under observation or awaiting attention by the human monitor. Since instruments are presumed to possess distinct characteristics, $i_j=0$ or 1 for $j=1,2,\ldots m.$. Thus a suitable particularized form of model solution taken from Posner and Bernholtz yields.

$$P(i_1,\ldots\ldots,i_m) = K(i_1+i_2+\ldots+i_m)!\, \rho_1^{i_1}\, \rho_2^{i_2}\ldots\rho_m^{i_m} \tag{7}$$

where the normalizing constant K is obtained by summing $P(i_1,\ldots,i_m)$ over alle possible combinations of $i_1\ldots,i_m$ satisfying $0 \leq i_1 + i_2 +\ldots+ i_m \leq m$, and where

$$\rho_j = (\mu_j U_j)^{-1},\ j = 1,2,\ldots,m$$

is the ratio of mean attention time to mean interdemand time for i_j .

For $i_1 + i_2 +\ldots+ i_m = n \leq m$, let P_n denote the stationary probability that n instruments are under observation or awaiting attention. Hence P_n is obtained as the summation of the $P(i_1,\ldots i_m)$ over all possible values of $i_1,\ldots,i_m$ such that $i_1 + i \ldots+i_m= n.$ Carrying out this summation operation on (7) yields

$$P_n = K\, n\,!\, S_{m,n} \tag{8}$$

where $S_{m,n}$ is generated using the following recursive procedure. If

$$S_{r,j} = \sum_{i_1 + i_2 + \dots + i_r = j} \rho_1^{i_1} \rho_2^{i_2} \dots \rho_r^{i_r}, \quad 1 \le j \le r \le m$$

then in general

$$S_{r,j} = S_{r-1,j} + \rho_r S_{r-1,\, j-1} \tag{9}$$

with

$$S_{r,o} = 1 \quad , \; r = 1,2,\dots,m$$

$$S_{r,1} = \sum_{n=1}^{r} \rho_n, \; r = 1,2,\dots,m$$

Furthermore, since $P_o + P_1 + \dots + P_m = 1$, the normalizing constant K is determined from (8) as

$$K^{-1} = \sum_{j=o}^{m} j!\, S_{m,j}$$

so that a closed form solution for P_n is

$$P_n = \frac{n!\, S_{m,n}}{\sum_{j=o}^{m} j!\, S_{m,j}}, \quad n = o,1,\dots,m \tag{10}$$

Thus, $P_o = K$ is the proportion of time the observer is idle, and

$$P_o + P_1 = K(1+\rho_1+\dots+\rho_m)$$

is the proportion of time that no instruments are queueing to await the observer's attention.

4.a Waiting Time to Observation for I_m

Suppose we are concerned with how long a particular instrument m (say) must wait in the queue before the human monitor is prepared to make his observation on that instrument. As noted, this (random) time is important since the plant under control of the human monitor might go out of limits during this waiting period of non-observation.

Let the random variable W_m denote the (stationary) waiting time in the queue for i_m. Furthermore, let $\vec{N}_m^{(r)} = (i_1,\dots,i_{m-1}) \equiv \vec{I}$ denote a particular subset of exactly r other instruments already in queue or in service at the moment of arrival of i_m. Thus, for example, if $i_j=0$ in this vector, then i_j is not present in the queue or in service when arrives, while $i_j=1$ indicates that i_j is present when i_m arrives.

Now W_m has a mixed probability distribution with the probability of no wait given by

$$\Pr\{W_m = o\} = P_o / \sum_{r=o}^{m-1} \sum_{i_1+i_2+\dots+i_{m-1}=r} \cdots \sum P(i_1, i_2, \dots, i_{m-1}, o) \tag{11}$$

and the density function associated with positive wait denoted by $f_{W_m}(\cdot)$ and distribution function given by $F_{W_m}(\cdot)$. Therefore

$$\Pr\{W_m > \tau\} = 1 - F_{W_m}(\tau) = \sum_{r=1}^{m-1} \sum_{\vec{i}=(i_1,\dots,i_{m-1})} \cdots \sum \Pr\{W_m > \tau \mid \vec{N}_m^{(r)} = \vec{i}\} \cdot \Pr\{\vec{N}_m^{(r)} = \vec{i}\} \tag{12}$$

where the multiple summation is carried out over all elements of $\vec{i}$ satisfying $i_1 + i_2 + \dots + i_{m-1} = r$.

Now, since $\Pr\{\vec{N}_m^{(r)} = \vec{i}\}$ is the conditional probability of a specific subset $\vec{i}$ of size r in the system at the moment of arrival of i_m we can therefore write.

$$\Pr\{\vec{N}_m^{(r)} = \vec{i}\} = P(i_1, i_2, \dots, i_{m-1}, o) / \sum_{k=o}^{m-1} \sum_{j_1+\dots+j_{m-1}=k} \cdots \sum P(j_1, \dots, j_{m-1}, o) \tag{13}$$

where $j_\gamma = o$ or 1 for $\gamma = 1, 2, \dots m-1$. From equation (7) and (8) we then have

$$\Pr\{\vec{N}_m^{(r)} = \vec{i}\} = r!\, \rho_1^{i_1} \cdots \rho_{m-1}^{i_{m-1}} / \sum_{n=o}^{m-1} n!\, S_{m-1,n} \tag{14}$$

Denote the service time of i_k $(K = 1, 2, \dots, m-1)$ by

$$G_{i_k}^k = \begin{cases} G_k & , i_k = 1 \\ o & , i_k = o \end{cases}$$

Then $\Pr\{W_m > \tau \mid \vec{N}_m^{(r)} = \vec{i}\} = \Pr\{G_{i_1}^1 + G_{i_2}^2 + \dots G_{i_{m-1}}^{m-1} > \tau\}$

Since, for $i_k = 1, G_k$ is exponentially distributed with parameter μ_k and since all service times represent mutually independent random variables, it can easily be shown that

$$\Pr\{W_m > \tau \mid \vec{N}_m^{(r)} = \vec{i}\} = \sum_{\gamma=1}^{m-1} \delta_{i_\gamma, 1} \prod_{\substack{k=1 \\ k \neq \gamma}}^{m-1} \left(\frac{\mu_k}{\mu_k - \mu_\gamma} \right)^{i_k} e^{-\mu_\gamma \tau} \tag{15}$$

where $\delta_{i_\gamma 1} = 1$ if $i_\gamma = 1$ and is 0 otherwise. Combining (14) and (15) into (12) yields

$$\Pr\{W_m>\tau\} = \sum_{r=1}^{m-1} r! \sum_{i_1+\dots+i_{m-1}=r} \cdots \sum \sum_{\gamma=1}^{m-1} \delta_{i_\gamma,1}\, \rho_\gamma \cdot$$

$$\prod_{\substack{k=1\\k\neq\gamma}}^{m-1} \left(\frac{\mu_k\, \rho_k}{\mu_k - \mu_k}\right)^{i_k} e^{-\mu_\gamma \tau} \Big/ \sum_{n=o}^{m-1} n!\, S_{m-1,n} \tag{16}$$

$$= \sum_{r=1}^{m-1} \sum_{r=1}^{m-1} \bar{T}^{\gamma}_{r,m-1}\; e^{-\mu_\gamma \tau}$$

where

$$\bar{T}^{\gamma}_{r,m-1} = r!\, \rho_\gamma\, T^{\gamma}_{r-1,m-1} \Big/ \sum_{r=o}^{m-1} n!\, S_{m-1,n}$$

$$T^{\gamma}_{r-1,m-1} = \sum_{i_1+i_2+\dots+i_{m-1}=r} \cdots \sum \delta_{i_\gamma,1} \prod_{\substack{k=1\\k\neq\gamma}}^{m-1} \phi_{k,\gamma}^{\,i_k}$$

$$= \sum_{i_1+\dots+i_{\gamma-1}+i_{\gamma+1}+\dots+i_{m-1}=r-1} \cdots \sum \prod_{\substack{k=1\\k\neq\gamma}}^{m-1} \phi_{k,\gamma}^{\,i_k} \tag{17}$$

$$\phi_{k,\gamma} = [(\mu_k-\mu_\gamma)\, U_k]^{-1}$$

Thus, the form of (16) reveals that the waiting time for i_m has a mixed exponential form involving the service parameters $\mu_\gamma(\gamma= 1,2,\dots,m-1)$ of all other instruments. The mean and variance of its waiting time to observation are therefore given by

$$E\,(W_m) = \sum_{r=1}^{m-1} \sum_{\gamma=1}^{m-1} \bar{T}^{\gamma}_{r,m-1} / \mu_\gamma$$

$$V\,(W_m) = \sum_{r=1}^{m-1} \sum_{\gamma=1}^{m-1} \bar{T}^{\gamma}_{r,m-1} / \mu_\gamma^{\,2}$$

In order to obtain analogous expressions for any other particular instruments, the method of this section is merely repeated using another instrument in the assumed m^{th} position. The mathematical formulations are symmetric in form so that implementation of these calculations is easily realized. In particular, to assist in creating similar constructs the formula for $T^{\gamma}_{r-1,m-1}$ in (17) may be obtained in a simplified recursive manner. By successive partitioning of the multiple summation and product form, it can be shown in general that

$$T^{\gamma}_{j,n} = T^{\gamma}_{j,n-1} + \phi_{n,\gamma}\, T^{\gamma}_{j-1,n-1}, \quad 1\le j\le n\le m-1,\ \gamma = 1,2,\dots,m-1$$

with

$$(1)\ T^{\gamma}_{j,\gamma} = T^{\gamma}_{j,\gamma-1}$$

$$(2)\ T^{\gamma}_{j,j} = \prod_{k=1}^{j} \phi_{k,\gamma}\,,\quad j=1,2,\ldots,\gamma-1$$

$$(3)\ T^{\gamma}_{j,j+1} = \prod_{\substack{k=1\\ k\neq\gamma}}^{j+1} \phi_{k,\gamma}\,,\quad j=\gamma,\gamma+1,\ldots,m-2$$

$$(4)\ T^{\gamma}_{o,n} = 1$$

4b. Probability of Plant Failure

Suppose that historical plant statistics indicate that the time τ_m required for instrument m to go out of limits measured from the time it is ready to be observed, is a random variable having some distribution function $F_{\tau_m}(\tau)$. Then the probability P_f^m that i_m will go out of limits during a wait for observation is given by

$$\begin{aligned} P_f^m &= \int_{\tau=o}^{\infty} Pr\,\{W_m > \tau\}\, dF_{\tau_m}(\tau) \\ &= \sum_{r=1}^{m-1} \sum_{\gamma=1}^{m-1} T^{\gamma}_{r,m-1} \int_{\tau=o}^{\infty} e^{-\mu_\gamma \tau}\, dF_{\tau_m}(\tau) \\ &= \sum_{r=1}^{m-1} \sum_{\gamma=1}^{m-1} T^{\gamma}_{r,m-1}\, L_m(\mu_\gamma) \end{aligned}$$

where

$$L_m\{s\} = \int_{o}^{\infty} e^{-s\tau}\, dF_{\tau_m}(\tau)\,,\quad Re\,\{s\} > o$$

is the Laplace-Stieltjes transform of the time to plant failure due to instrument m. Thus, for example if τ_m is known to be approximately normally distributed with mean α_m and variance σ_m^2 then

$$P_f^m = \sum_{r=1}^{m-1} \sum_{\gamma=1}^{m-1} T^{\gamma}_{r,m-1} \exp[-\mu_\gamma \alpha_m + \tau_m^2 \mu_\gamma^2 / 2] \tag{18}$$

As previously noted, analogous failure probability terms can be determined for other instruments under consideration. Denote these by $P_f^1, \ldots, P_f^{m-1}$. Since these failure probabilities are assumed to be quite small, failures due to each instrument going out of limits may be treated (approximately) independently so that a good estimate of plant failure is given by

$$P_f = \sum_{r=1}^{m} P_f^r \tag{19}$$

4.c Number of Plant Failures in Time V

Consider a reasonably long period V. How many plant failures are anticipated during this period on the basis of assumed observer behaviour? We have determined that i_m will go out of limits during a wait for the observer with probability P_f^m. Thus, on each cycle through the observer system instrument m will go out of limits with that probability. The mean cycle time for i_m is given by $U_m + E(W_m) + 1/\mu_m$ so that the expected number of failures during period V due to instrument m is

$$\frac{\mu_m \, V \, P_f^m}{1 + \mu_m U_m + \mu_m E(W_m)} \tag{20}$$

By treating instrument failures independently, the expected total number of plant failures is given by

$$\sum_{r=1}^{m} \frac{\mu_r P_f^r}{1 + \mu_r U_r + \mu_r E(W_r)} V \tag{21}$$

4.d General Comment

At this point the question of how to improve the reliability of the system in terms of reducing the expected plant failure rate to an acceptable level can be addressed. It is clear that it would be possible to apportion the care of instruments among several human monitors so that each would be responsible for a specific subset of the instrument array. From the analysis presented, subsystem reliablity corresponding to a subset of instruments under the care of a single operator can be evaluated, leading to a reliability measure for the entire system with several operators working simultaneously. An iterative procedure can then easily be developed for allocating instruments to operators in order to further improve the total system reliability. This allocation procedure could continue until an acceptable performance level is obtained, or possibly until an optimal allocation results. The analysis serves, therefore, to provide solutions both to the problems of reliability and to the problem of manning requirements. Further applications of the results to questions of training, simulation, and operator evaluation are readily seen.

5. References

I. J.W. Senders, J.E. Elkind, M.C. Grignetti, and R.P. Smallwood, An Investigation of the Visual Sampling Behaviour of Human Observers, NASA Cr 434, (1965).

APPENDIX E

A Re-Analysis of the Pilot Eye-Movement Data

INTRODUCTION AND BACKGROUND

In an earlier paper (1) we have considered the question of the distribution of attention among a number of information sources. There we started with some primitive considerations about the behaviour of human observers and arrived at a theoretical description of that behaviour which predicted quite well the manner in which observers monitor instruments when faced with the task of monitoring many independent instruments. Our basic assumptions were as follows : 1) that an operator samples an instrument at a rate which is proportional to the effective bandwidth of the signal presented on that instrument; 2) that an operator has a constant input channel capacity (which implies) 3) the duration of his observation will, on the average, be proportional to the average amount of information presented in each look at each instrument; 4) that the probability of transition from one instrument to another is simply the product of the probability of being on the first instrument and the probability of being on the second instrument (which implies), 5) that an operator can make a transition from an instrument to itself.

The results of the analysis, when applied to real operators observing a set of four instruments presenting random functions of various bandwidths, are shown in Table 1 and Fig. 1. In general, it would appear that operators faced with the task of monitoring four independent time functions behaved like simple zero-order Markov processors subject to the constraints of sampling frequencies and input channel capacities.

A more difficult problem of analysis is presented by the task of sampling many instruments which are part of a coordinated system like an aircraft with its instrument, sensors, and couplings. A number of years ago a series of studies done at Wright Air Development Center was aimed at finding a solution to the question of instrument panel layout in military aircraft (2)-(10). Presumably, if an instrument is looked at frequently, it follows that it should be placed in some central or prominent position. Further, if two instruments are looked at in succession more frequently than either is looked at in conjunction with any other instrument, it follows that they should be placed in adjacent locations. The investigators, by recording eye-fixations of pilots in a series of flight tests, were able to determine how often and for how long the various instruments were looked at in a number of flight conditions. They also identified the pairs of instruments looked at in successive looks.

Unfortunately, no attempt was made to record the indications presented on the instruments which were being fixated. As a result, it has been impossible to use the data to test for a

TABLE I

CALCULATED AND MEASURED TRANSITION PROBABILITIES*

Transition	Calculated Transition Probabilities	Measured Transition Probabilities
P 4.2 = (3.17) (0.537)	(0.214) = 0.364	0.324
P 4.1 = (3.17) (0.537)	(0.172) = 0.293	0.297
P 4.1/2 = (3.17) (0.537)	(0.077) = 0.131	0.133
P 2.1 = (3.17) (0.214)	(0.172) = 0.117	0.112
P 2.1/2 = (3.17) (0.214)	(0.077) = 0.052	0.051
P 1.1/2 = (3.17) (0.172)	(0.077) = 0.042	0.040
	0.999	0.957

* Adapted from Senders [1].
4.3 percent of the eye fixations were to "other" places in the visual field.

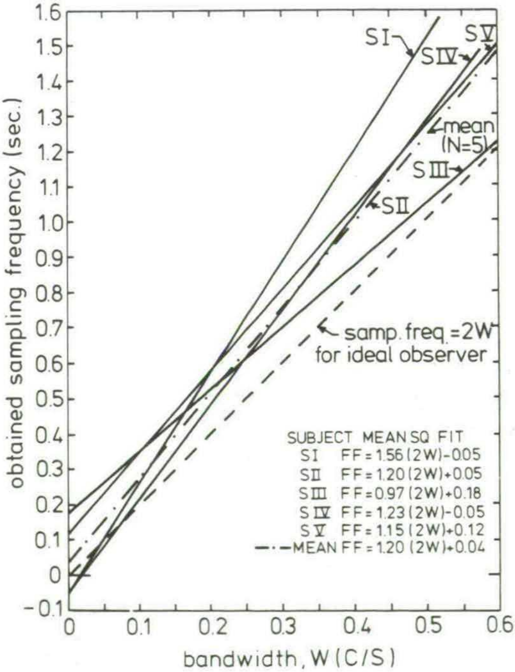


Fig. 1. Regression of obtained sampling frequencies on bandwidth.

relationship between the signal being presented on the instrument and the frequency and duration of observations of it. It is possible, however, to apply certain other parts of our model to the data obtained in those studies. We can ask whether our prediction of "link values", or two-way transition probalilities, based purely upon the state probability of each of the instruments, will be a good match for the link values recorded in flight. If, on the basis of state probabilities, we can reasonably well predict the link values , a number of interesting conclusions can be reached. In the first place, we can establish the degree to which the calculated transition probabilities would have served as the basis for instrument panel design and layout. Secondly, by comparing the predicted and measured values and noting the major discrepancies, we may be led to some degree of understanding of the orderly scanning process, if one exists. A large discrepancy between the predicted and measured values for a pair of instruments suggests that a scanning pattern relating those two instruments did, in fact, exist.

Data

As available information we have records of the frequency and duration of observations of each of the instruments in the cockpit during various flight maneuvers. If we were to sample the monitoring process at any point in time, the probability that we would find the pilot fixating on instrument *A* is the product of the frequency and the duration of observation of instrument *A*. (For the moment and for simplicity, we will assume that transitions are instantaneous).

$$P_a = F_a \cdot D_a \qquad \begin{aligned} F &= \text{frequency of observation} \\ D &= \text{duration of absevation} \end{aligned} \tag{1}$$

TABLE II

CALCULATED AND MEASURED TRANSITION PROBABILITIES FOR NIGHT GCA, STUDY III

Instrument	Proportion of Time Spent on Each Instrument	Two-Way Transitions	Percent of the Total Predicted Link Values	Percent of the Total Observed Link Values
AS	0.17	AS-DG	0.2501	0.2875
		AS-GH	0.0951	0.0595
DG	0.50	AS-ALT	0.0150	0.0354
		AS-TB	0.0100	0.0100
GH	0.19	AS-VS	0.0250	0.0163
		AS-EI	0.0200	0.0063
ALT	0.03			
		DG-GH	0.2796	0.3064
TB	0.02	DG-ALT	0.0441	0.0396
		DG-TB	0.0294	0.0297
VS	0.05	DG-VS	0.0736	0.0484
		DG-EI	0.0589	0.0272
		GH-ALT	0.0168	0.0126
		GH-TB	0.0119	0.0198
		GH-VS	0.0280	0.0467
		GH-EI	0.0224	0.0180
		ALT-TB	0.0018	0.0088
Total Number of Links = 7382		ALT-VS	0.0044	0.0130
		ALT-EI	0.0035	0.0004
		TB-VS	0.0029	0.0042
		TB-EI	0.0024	0.0019
		VS-EI	0.0059	0.0077
			$\sum$ 1.0008	$\sum$ 0.9975

TABLE III

CALCULATED AND MEASURED TRANSITION PROBABILITIES FOR STRAIGHT AND LEVEL FLIGHT, STUDY V

Instrument	Proportion of Time Spent on Each Instrument	Two-Way Transitions	Percent of the Total Predicted Link Values	Percent of the Total Observed Link Values
AS	0.08	AS–DG	0.0868	0.0730
		AS–GH	0.0586	0.0300
DG	0.40	AS–ALT	0.0304	0.0530
		AS–TB	0.0065	0.0020
GH	0.27	AS–VS	0.0109	0.0070
		AS–EI	0.0043	0.0010
ALT	0.14			
		DG–GH	0.2930	0.3390
TB	0.03	DG–ALT	0.1519	0.1610
		DG–TB	0.0326	0.0250
VS	0.05	DG–VS	0.0543	0.0190
		DG–EI	0.0217	0.0060
EI	0.02			
		GH–ALT	0.1025	0.0650
		GH–TB	0.0270	0.0270
		GH–VS	0.0366	0.0780
		GH–EI	0.0146	0.0130
		ALT–TB	0.0114	0.0370
Total Number of Links = 1548		ALT–VS	0.0190	0.0430
		ALT–EI	0.0076	0.0000
		TB–VS	0.0041	0.0210
		TB–EI	0.0043	0.0000
		VS–EI	0.0027	0.0020
			$\sum$ 0.9758	$\sum$ 1.0020

TABLE IV

CALCULATED AND MEASURED TRANSITION PROBABILITIES FOR NIGHT GCA, STUDY VII

Instrument	Proportion of Time Spent on Each Instrument	Two-Way Transitions	Percent of the Total Predicted Link Values	Percent of the Total Observed Link Values
AS	0.15	AS–XPT	0.0044	0.0198
		AS–GH	0.0348	0.0617
XPT	0.01	AS–EI	0.0087	0.0035
		AS–ALT	0.0174	0.0321
GH	0.08	AS–DG	0.2209	0.1773
		AS–VS	0.0871	0.0538
EI	0.02	AS–TB	0.0000	0.0005
ALT	0.04	XPT–GH	0.0023	0.0193
		XPT–EI	0.0006	0.0015
DG	0.50	XPT–ALT	0.0012	0.0010
		XPT–DG	0.0147	0.0199
VS	0.20	XPT–VS	0.0058	0.0025
		XPT–TB	0.0000	0.0000
TB	0			
		GH–EI	0.0046	0.0035
		GH–ALT	0.0093	0.0119
		GH–DG	0.1178	0.0775
		GH–VS	0.0464	0.0859
		GH–TB	0.0000	0.0000
		EI–ALT	0.0023	0.0064
		EI–DG	0.0295	0.0119
Total Number of Links = 2035		EI–VS	0.0116	0.0099
		EI–TB	0.0000	0.0000
		ALT–DG	0.0589	0.0741
		ALT–VS	0.0232	0.0291
		ALT–TB	0.0000	0.0000
		DG–VS	0.2946	0.3067
		DG–TB	0.0000	0.0000
		VS–TB	0.0000	0.0005
			$\sum$ 0.9961	$\sum$ 1.010

Let us briefly review the method used to obtain the basic information to which we will apply our Markovian model. Various numbers of Air Force pilots flew a variety of propeller and jet aircraft (2) - (10). A motion picture camera was mounted behind the pilot's seat and a mirror mounted on the instrument panel in such a way that the camera viewed the pilot's head and eyes in the mirror. Calibration runs were made in which the pilots successively fixated on each of the instruments on the instrument panel in order to provide reference photographs for the film readers. The films taken in flight were later compared with the standard calibration fixations by at least two independent sets of readers, and each frame of film was identified as to its probable fixation point. Since the frames were sequentially arranged, it was also possible to record the previous fixation. This laborious task has perhaps been outmoded by more elegant means of recording eye position, but the data which were

TABLE V

CALCULATED AND MEASURED TRANSITION PROBABILITIES FOR DAY STRAIGHT AND LEVEL, STUDY IX

Instrument	Proportion of Time Spent on Each Instrument	Two-Way Transitions	Percent of the Total Predicted Link Values	Percent of the Total Observed Link Values
AS	0.12	AS-XPT	0.0060	0.0182
		AS-GH	0.0508	0.0575
XPT	0.02	AS-EI	0.0149	0.0079
		AS-ALT	0.0598	0.0957
GH	0.17	AS-DG	0.0926	0.0382
		AS-VS	0.0359	0.0133
EI	0.05	AS-TB	0.0030	0.0000
ALT	0.20	XPT-GH	0.0085	0.0309
		XPT-EI	0.0025	0.0018
DG	0.31	XPT-ALT	0.0100	0.0012
		XPT-DG	0.0134	0.0218
VS	0.12	XPT-VS	0.0060	0.0042
		XPT-TB	0.0005	0.0006
TB	0.01			
		GH-EI	0.0212	0.0145
		GH-ALT	0.0847	0.0697
		GH-DG	0.1312	0.0812
		GH-VS	0.0508	0.0715
		GH-TB	0.0042	0.0042
		EI-ALT	0.0249	0.0218
		EI-DG	0.0386	0.0073
Total Number of Links = 1651		EI-VS	0.0149	0.0055
		EI-TB	0.0012	0.0012
		ALT-DG	0.1544	0.1993
		ALT-VS	0.0598	0.0636
		ALT-TB	0.0050	0.0000
		DG-VS	0.0926	0.1599
		DG-TB	0.0077	0.0036
		VS-TB	0.0030	0.0055
			$\sum$ 1.0000	$\sum$ 1.0000

obtained apparently had good reliability and are good indexes of the frequency and duration of observation for the various instruments.

CALCULATIONS AND RESULTS

We suppose that the probability of being on instrument A is P_a and the probability of being on instrument B is P_b. What then is the probability that a transition will occur from instrument A to instrument B ? Since we assume that the probability of a transition to any instrument is merely the probability of being on that instrument, the transition probability (in one direction) is the product of the two state probabilities $P_{\overline{ab}} = P_a P_b$ and the probability of a transition in either direction is $2P_a P_b$. However, since the eye may make a "transition" from any instrument to that same instrument, the number of transitions which can be recorded by the cameras is less than the total number of transitions which takes place. The "repeated looks" are indistinguishable from "longer looks" which we also suppose to occur. The observable transitions are those between different instruments. As a result the probability of observing a transition between instruments A and B will be greater than the calculated probability $2P_{\overline{ab}}$. Since a pair of looks on any instrument i has a probability P_i^2, the measured frequency of transitions between A and B will be

$$P_{\overline{ab}} \text{ meas.} = \frac{2P_{\overline{ab}}}{1 - \sum_i P^2_{(i)}} \tag{2}$$

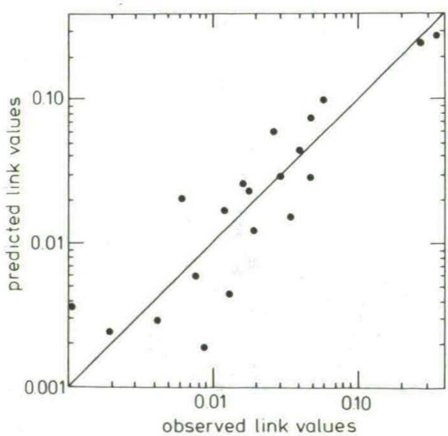


Fig. 2. Probabilities for night GCA, Study III. The points are the predicted percentages for the various transitions plotted against the obtained percentages based on 7382 transitions.

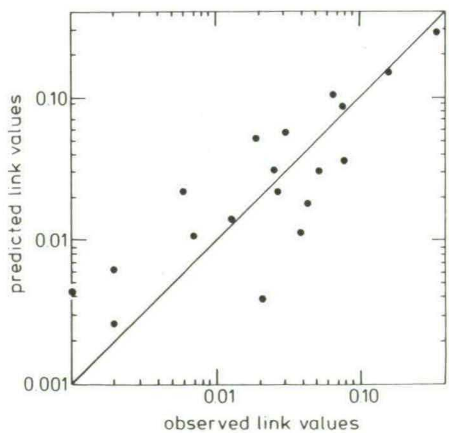
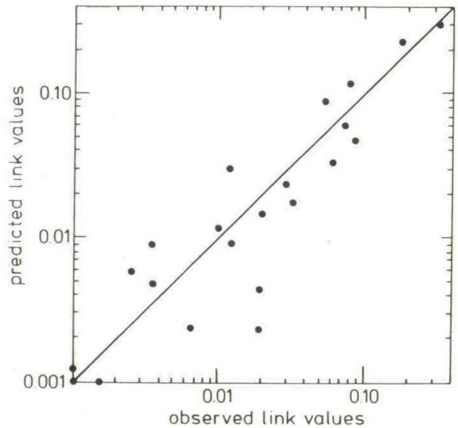


Fig. 3. Probabilities for straight and level flight, Study V. The points are the predicted percentages for the various transitions plotted against the obtained percentages based on 1548 transitions.



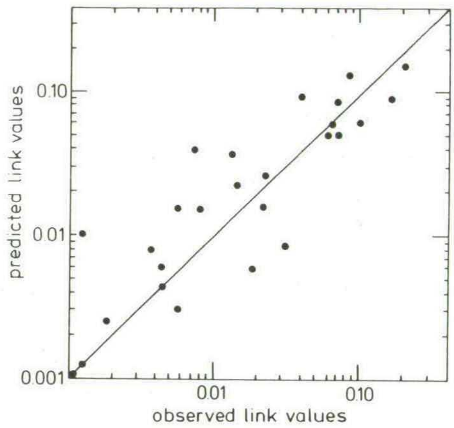


Fig. 4. Probabilities for night GCA, Study VII. The points are the predicted percentages for the various transitions plotted against the obtained percentages based on 2035 transitions.

Fig. 5. Probabilities for day straight and level, Study IX. The points are the predicted percentages for the various transitions plotted against the obtained percentages based on 1651 transitions.

The calculations have been carried out for the data from the flight tests and are presented in Tables II-V and Figs. 2-5.

DISCUSSION

The general relationship between the predicted and observed values is strong. In particular, it would have been possible to arrange the instruments on the basis of the calculated transition probabilities about as well as on the basis of the measured link values. (The large percentage differences at the lower values are of no great significance).

Why would there be any error in prediction at all ? There are a number of plausible explanations for the differences. Probably the most obvious and important one is that pilots in flying aircraft do in fact utilize regular and ordered scanning patterns in a way dictated by the coupling in the system or by the kind of information which is being presented. Thus, it seems reasonable to assume that if, on observation of the altimeter, an error is observed between that which is displayed and that which is desired, the next observation should be on the vertical speed indicator in order to observe the cause of the change in altitude. If there is a deviation in the vertical speed indicator, the next observation might well be on the gyro-horizon since it could be a change in attitude which led to a change in rate-of-climb which, in turn, led to the error in altitude. Secondly, there is the possibility that despite the relative efficiency of the zero-order Markov model which we have considered, the pilot in fact operates as a higher-order Markov process and has a probability of a transition to another instrument which depends on the instrument currently being observed. Lastly, there is also a high probability that the operator reading an instrument takes note not only of the position of indicator but also of its velocity. If he behaves as a rational sampling system dealing with position and velocity, and having an error tolerance band, he may modify his sequence of observations on the basis of the sign and magnitude of the velocity and the closeness of the indication to the limits of the acceptable band.

Examination of the results shows that there is good agreement between prediction and observation for the higher values. Since the data are plotted logarithmically, the deviation of the points from the 45-degree line are indicative of percentage deviation rather than of the absolute amount.

If we examine in a more intuitive manner the implications of (1) and (2), we are led to the following chain of reasoning :

1) Due to the requirement of other instruments for fairly frequent sampling, no instrument can be looked at for a very long time.

2) As a result, those instruments which have high probabilities have them in consequence of having been sampled many times for a relatively short duration rather than infrequently for long durations.

3) Thus, the instruments with high probabilities must also have a large number of arrivals and departures and, therefore, a large number of observable transitions.

4) Similarly, since there is some minimum time which can be spent in switching to an instrument and making the least observation of it, those instruments which have low probabilities must have a small number of arrivals and departures. (Of course, ʼnat these two simple conclusions are in fact true is evident from the flight data. What is to the point is that they *must* be true).

5) Therefore, large numbers of transitions cannot take place between instruments which have small numbers of arrivals and departures, and most transitions must involve those pairs of instruments both members of which have large numbers of arrivals and departures

Since the space available on an instrument panel is limited, and since the most desirable areas are already identified, the design and layout of the panel will depend largely on the high-probability instruments, with the others having only small freedom of location and relative position. From the practical point of view, therefore, although the flight data are significantly different from the predicted results, the design process would probably not be affected. If even as fair as relationship could be shown to exist between the characteristics of the signal presented on an instrument and the frequencies and durations of observation, then it follows that the design could have been based purely on an engineering analysis of the physical characteristics of the aircraft, its sensors, and its physical environment, supplemented by some specification of the mission of the reading accuracy demanded. That such information may not *now be* readily available is only a temporary obstacle to a rational design procedure.

CONCLUSIONS

Since records of the instrument reading during the flights were not taken, there has been no opportunity afforded by these data to test the predictions of theory as to frequency of sampling of the various instruments and total time spent on each instrument. Given the measured time spent on each instrument, it has been possible to test the transition model. The results of this show that the transition model does not predict within reasonable error (from the theoretical point of view) the "Link Values" observed in flight. However, the constraints on transitions imposed by frequency of sampling enforce a general adherence to the model. As a result, althought it is almost surely the case that scanning patterns exist, they must coexist with the demands of airthmetic.

For design purposes, the predictions of the model appear good enough. Where there exist three or more instruments with very similar probabilities of sampling, there is a greater freedom for the observer to use scanning patterns. Therefore, when three or more such instruments exist in a system, more caution must be used in applying the model. Where only one or two instruments dominate the array, far less freedom can exist. In future flight tests or simulator studies of eye movements and sampling, complete records should be taken of all the signals entering and leaving the cockpit if a complete understanding is desired of the work that the pilot is doing.

REFERENCES

(1) J.W. Senders, "The human operator as a monitor and controller of multidegree of freedom systems" *"IRE Trans. on Human Factors in Electronics,* vol. HFE-5, pp. 2-5. September 1964.

(2) R.E. Jones, J.L. Milton, and P.M. Fitts, "Eye fixations of aircraft pilots : A review of prior-eye-movement studies and a description of a technique for recording the frequency, duration, and sequence of eye fixations during instrument flight" USAF Tech, Rept. 5837 (ATI 65996). September 1949.

(3) J.L. Milton, R.E. Jones, and P.M. Fitts, "Eye fixations of aircraft pilots : Frequency, duration and sequence of fixations when flying the USAF instrument low approach systems (ILAS)," UASF Tech. Rept. 5839 (ATI 65996), October 1949.

(4) P.M. Fitts, R.E. Jones, and J.L. Milton "Eye fixations of aircraft pilots : Frequency, duration, and sequence of fixations when flying Air Force ground controlled approach system (GCA)," AF Tech. Rept. 5967 (ATI 69038) February 1950.

(5) R.E. Jones, J.L. Milton, and P.M. Fitts, "Eye fixations of aircraft pilots : Frequency, duration and sequence of fixations during routine instrument flight" AF Tech. Rept. 5975 (ATI 73422) December 1949.

(6) J.L. Milton, R.E. Jones, and P.M. Fitts, "Eye fixations of aircraft pilots : Frequency duration and sequence of fixations when flying selected maneuvers during instrument and visual flight conditions" AF Tech. Rept. 6018 (ATI 84010) August 1950.

(7) J.L. Milton, B.B. McIntosh and E.L. Cole, "Eye fixations of aircraft pilots : Fixations during day and night ILAS approaches using an experimental instrument panel arrangement," USAF Tech. Rept. 6570 (ATI 128350), October 1951.

(8) J.L. Milton, B.B. McIntosh, and E.L. Cole "Eye fixations on aircraft pilots : Fixations during day and night GCA approaches using an experimental instrument panel arrangement," USAF Tech. Rept. 6709 (ATI 158203), February 1952.

(9) J.L. Milton and F.J. Wolfe "Eye fixations of aircraft pilots : Fixations during zero reader approaches in a jet aircraft," WADC Tech. Rept. 52-17 (ATI 149050, February 1952.

(10) J.L. Milton, B.B. McIntosh, and E.L. Cole "Eye fixations of aircraft pilots : Routine maneuvers under day and night conditions, using an experimental panel arrangement," WADC Tech. Rept. 53220 (AD 35161) March 1954.

ABSTRACT

The visual scanning problem derives directly from a wide variety of real world tasks. The specific task of the aircraft pilot as the monitor of a large number of state variables is described and the extensive studies of the pilot's eye movements summarized and criticized. The pilot's task is used as the model for a general theory of scanning based on the notion that scanning is a sampling process which in turn generates a queueing process.

A hierachy of mathematical models of scanning is derived from the general theory. The models are based on a number of possible strategies assumed to drive the behaviour of the instrument monitor. The simplest (and most unrealistic) model is based on the assumption that the scanner is a periodic sampler attempting, on the basis of the samples, to reconstruct the time function displayed on the instrument. The next model assumes that the choice of instrument next to be fixated is random with the selection probability of each instrument equal to the proportion of the total information flow presented by each. Then follow various assumptions about conditional sampling. This gives rise to models based on the realistic notion that what is seen in one look at an instrument is the basis for the choice of when to make the next look. The interval is computed on the basis of an assumed goal. One goal is to be looking at the instrument when the probability that the displayed signal has been exceeded a limit is maximum. Another is to be looking when that probability has itself exceeded some threshold. Another model assumes that the relative precision of reading of the displayed variable (and therefore the information flow rate) is a function of how close to the limit the variable is. The effective reduction of signal bandwidth results, when the displayed value is far from the limit, in a longer interval. Finally a model based on rate detection and imperfect memory is presented. Each of these model derived from the general theory yields a different relation between fixation frequency and signal bandwidth.

The models also make it possible to predict the probability of shift of fixation from any instrument to any other, on the assumption of independence of selection. A model of sampling of globally presented information, as in automobile driving, is presented in an appendix as is an analysis of the queueing which results from the instrument monitoring task.

Another appendix presents a model of workload based on an internal (covert) sampling of the elementary components of a complex task. A final appendix compares Fitts' data on Link Values with the predictions of the random choice model.

Experiments were carried out on arrays of four and six instruments. Eye movements of practiced scanners were recorded and the intervals and durations of fixation on each of the displays calculated. The experimental data allow only partial test of the various models. The general theory is strongly supported, but it is difficult to discriminate among the various models on the basis of the data. One or more of the model strategies may be simultaneously operating in any subject. The predictions of the Link Value model are well supported by the experimental data as well as by the data from the pilot eye movement studies of Fitts et al.

The theory is of importance in both applied and theoretical ways. It assists in the calculation of workloads, manning requirements, and man-machine system reliabilities in complex tasks like those of the control room of a nuclear power plant or a large aircraft. The theory has implications for the problem of attention especially with regard to the notions of covert internal scanning processes. It is as much a theory of attention as of the specific manifestation of attention demonstrated by scanning.

SAMENVATTING

Vragen met betrekking tor visueel aftasten van de omgeving worden in een grote diversiteit van praktijktaken aangetroffen. De specifieke taak van een piloot in het bewaken van de toestand van een groot aantal instrumenten wordt beschreven. De omvangrijke studies voor oogbewegingen van piloten worden samengevat en becritiseerd. De taak van een piloot wordt als model gebruikt voor een algemene theorie over visueel aftasten, gebaseerd op het idee dat aftasten bemonstering van informatie impliceert. Dit roept op zijn beurt een volgorde probleem van bemonstering op.

Op deze algemene theorie wordt een hiërarchie van mathematische modellen gebouwd. Deze modellen zijn gebaseerd op mogelijke strategieën die door de waarnemer kunnen worden gehanteerd. Het eenvoudigste, maar tegelijk wat onwerkelijke model betreft de aanname dat de waarnemer periodiek de informatiebronnen bemonstert om op grond hiervan de tijdfunktie van het instrument te reconstrueren. Het volgende model neemt aan dat de keuze van fixatie van de diverse instrumenten willekeurig is, waarbij de waarschijnlijkheid van fixatie van een bepaald instrument gelijk is aan het percentage informatie dat door ieder instrument wordt gegenereerd. Vervolgens zijn er diverse aannamen met betrekking tot voorwaardelijk bemonsteren. Deze leiden tot modellen die uitgaan van het realistische idee dat wat op een bepaald moment op een instrument wordt gezien bepalend is voor het moment dat dit instrument opnieuw wordt bemonsterd. Het interval tussen successieve fixaties wordt berekend op grond van een veronderstelde doelstelling : Een doelstelling is het instrument te fixeren als de waarschijnlijkheid dat het signaal een bepaalde waarde overschreden heeft maximaal is. Een andere doelstelling is om te fixeren als die waarschijlijkheid een bepaalde drempel overschrijdt. Nog een ander model gaat ervan uit dat de relatieve leesnauwkeurigheid van de betreffende variable (en daarmee de mate van informatietransmissie) een funktie is van de afstand van de waarde van de variable tot zijn toelaatbare grenswaarde. De effectieve afname van de bandbreedte van het signaal leidt tot een langer interval tussen fixaties, naarmate de waarde verder van de grenswaarde verwijderd is. Tenslotte wordt een model ontwikkeld op grond van detektie van de mate van informatietransmissie en van onvolledig onthouden van de toestand van het signaal. Elk van deze modellen leidt tot een ander verband tussen frekwentie van fixatie van een instrument en de bandbreedte van het signal.

De modellen bieden ook de mogelijkheid om de waarschijnlijkheid van fixatieverschuiving van een instrument naar een ander instrument te voorspellen, uitgaande van onafhankelijkheid van selektie. Een model betreffende bemonstering van globaal aangeboden informatie, zoals bij autorijden, wordt in Appendix B ontwikkeld. Verder geeft Appendix D een analyse van rijvorming zoals die zich bij het bewaken van een groter aantal instrumenten kan voordoen. Appendix C behandelt een model over werkbelasting uitgaande van interne bemonstering van de elementaire componenten van een complex taak. Een laatste appendix vergelijkt Fitts' gegevens over verbanden tussen instrumenten met de voorspellingen van het willekeurige keuzemodel.

De experimenten betreffen het bewaken van een rij van vier of zes instrumenten. Oogbewegingen van geoefende waarnemers werden opgenomen. evenals de intervallen en fixatieduur voor elk instrument. Deze experimentele gegevens kunnen de verschillende modellen slechts gedeeltelijk onderzoeken. De algemene uitgangspunten van de theorie worden krachtig ondertseund maar het is toch moeilijk om de diverse modellen goed van elkaar te onderscheiden. Bovendien kunnen één of meer strategieën tegelijk door een proefpersoon worden gebruikt. De voorspellingen van het model over de onderlinge relaties tussen instrumenten worden goed bevestigd door de experimentele gegevens alsook door de oogbewegingsstudies van Fitts en zijn medewerkers.

De theorie is van belang voor toepassing en modelvorming. Zij maakt het mogelijk om werkbelasting en personeelsbezetting te berekenen alsook om de betrouwbaarheid te beoordelen van mens-machine systemen in complexe taken zoals men die aantreft in een groot vliegtuig of in de controlekamer van een kernenergiecentrale. Verder heeft de theorie implikaties voor het aandachtsbegrip, speciaal waar het intern aftasten van de omgeving betreft. De theorie is evenzeer een theorie over aandacht als over de specifieke manifestatie van aandacht zoals die zich uit in het aftasten van de omgeving.

APPENDIX A. The human operator as a monitor and controller of multidegree of freedom systems.
IEEE Transactions on Human Factors in Electronics, 1964, Vol. 5, 2-5.

APPENDIX B. The Attentional demand of automobile driving.
Highway Research Record, 1967, 195, 15-32.

APPENDIX C. Axiomatic models of workload.
In : N. Moray (Ed.) Mental workload : its theory and measurement. Plenum, New York.

APPENDIX D. A Queueing model of monitoring and supervisory behaviour.
In : Sheridan, T.B. and Johannsen, G. (eds.) Monitoring behaviour and supervisory control, 1976, Plenum, New York.

APPENDIX E. A Re-analysis of the Pilot Eye-movement data.
IEEE Transactions on Human Factors in Electronics, 1966, Vol. 7, 103-106.

Appendix B : *co-authors* : A.B. Kristofferson, W.H. Levison, C.W. Dietrich and J.L. Ward

Appendix D : *co-author* : M.J.M. Posner.

Drukkerij
Neo Print
SOEST